A Fuzzy K-Nearest Neighbour-based Model for Detecting Lameness in Cattle

*[1]Opeyemi O. Adanigbo and [2]Oyeyemi T. Oyewole
[1] Department of Computer Engineering, Federal University, Oye-Ekiti, Nigeria
[2]Department of Computer Engineering, Elizade University, Ilara-mokin, Nigeria
opeyemi.adanigbo@fuoye.edu.ng|temitayo.oyeyemi@elizadeuniversity.edu.ng

Received: 21-MAY-2023; Reviewed: 13-JUN-2023; Accepted: 14-JUN-2023
http://doi.org/10.46792/fuoyejet.v8i2.1048

ORIGINAL RESEARCH

Abstract- In Africa, cattle are reared for meat and milk production and lameness is considered a major problem in modern dairy farming. Several studies have attempted to explicate automatic lameness detection systems using different techniques. However, these detection techniques are easily impacted by the physiological attributes of individual cows, ensuing in imprecise lameness detection. Consequently, this study presents a description and assessment of the performance of a fuzzy k-nearest neighbor (FKNN)-based classification system for automatic lameness detection from sensor data with a view to improving cattle lameness detection accuracy by reducing the rates of false-positive alerts. In order to further improve the model detection accuracy, Principal Component Analysis (PCA) was used in projecting the features on to a lower dimension on which the optimal FKNN model was formulated. The proposed system was tested using Classification Accuracy (Acc), sensitivity, specificity and the area under the receiver operating characteristic curve (ROC) curve (AUC) as performance metrics.

Keywords- Fuzzy K-Nearest Neighbour, Lameness, Principal component analysis.

1 Introduction

Animal production is the second largest source of food for human use, after crop farming. Animal husbandry contributes about 50% of the global agricultural gross domestic product and accounts for the livelihood sustenance and of about 1.3 billion people in underdeveloped countries (Herrero *et al*., 2016). In order to rise up to meet human needs, the livestock industry needs to improve the efficiency of its operations in order to increase animal productivity.

Lameness is considered the most important significant precursor of many locomotive ailments in Cattle and as such is considered severe in livestock animals. Cattle lameness is any deviance in gait due to pain or discomfort from hoof or leg injuries and disease (Van Hertem *et al*., 2014). It is one of the top three cow health issues related to economic losses in the dairy industry (Afonso *et al*.,2020) due to poorer reproductive performance, loss of milk production (Sjöström *et al*., 2018). This makes early detection of cattle ailments desirable in order to amply boost cattle productivity, automation of livestock monitoring is required.

*Corresponding Author

Section B- ELECTRICAL/COMPUTER ENGINEERING & RELATED SCIENCES
Can be cited as:
Adanigbo O.O. and Oyewole O. T. (2023). A Fuzzy K-Nearest Neighbour-based model for detecting lameness in Cattle, FUOYE Journal of Engineering and Technology (FUOYEJET), 8(2), 167-171. http://.doi.org/10.46792/fuoyejet.v8i2.1048

2 Related Works

Manual scoring of movement is time consuming, subjective, expensive and labour intensive. This is not possible with large herd sizes. In contrast, the automated moving scoring system (ALSS) has clear potential to provide a non-subjective and coherent method of lameness assessment (Gardenier *et al*., 2021). These included the animal's step count (Chapinal *et al*., 2010), the difference between step count and gait, back flex-ion, and head movements on a 2, 3, 4, and 5-point scale.

Some studies have used two-dimensional (2D) computer vision to analyse gait. These studies are focused around the quantification of different gait and posture variables such as back arch curve, step overlap (Pluk *et al*., 2010), the angles at which the hoofs are bent (Pluk *et al*., 2012), body movement pattern (Viazzi *et al*., 2014), variations in steps and gait (Van Hertem *et al*., 2014), assessment of speed decrement (O'leary *et al*., 2020). Also, force platforms (Liu *et al*., 2011), two-dimensional (2D) and three-dimensional (3D) cameras (Viazzi *et al*., 2014), and 2D or 3D accelerometers/ sensors attached to animals' limbs (Thorup *et al*., 2015) are the usually used sensors for automatic detection of lame cattle. Measurements obtained from these sensors are fed into machine learning algorithms to calculate lameness attributes such as step overlap (Pluk *et al*., 2010).

Automatic lameness detection approaches can be classified into three major groups (Schlageter-Tello *et al*., (2014); Alsaaod *et al*., 2019), namely: kinetic (measuring forces involved in locomotion), kinematic (measuring limb trajectories in space and time, and related specific posture variables), and indirect measurement approaches (these take behavioural variables into record). In this study, a combi-nation of the feeding behaviour, standing time, lying time and overall animal motion index were considered.

3 Methodology

In this study, PCA and FKNN were employed for feature reduction, classification and prediction of lameness respectively. A synthetic dataset was randomly generated using a two-week data recorded at Obafemi Awolowo University teaching and research farms, Ile-Ife as a template. The data was segmented into three categories viz: sound, mildly lame, and severely lame with the class labels 1, 2, and 3 allocated to each of the three different categories respectively.

3.1 The Fuzzy k-Nearest Neighbour Technique

As an improved variant of the KNN (k-Nearest Neighbour algorithm (Duda and Hart, 1973)), the FKNN classifier (Keller *et al*., (1985)) is a classification algorithm which incorporates the fuzzy set theory into KNN (Keller *et al*., 1985). In FKNN, the fuzzy memberships of samples are assigned to different classes. The class assigned to be the winner is the one which has the highest membership degree. For pattern x, the FKNN algorithm attributes a membership vector as a function of the pattern's distance from its k-nearest neighbours. The FKNN algorithm first computes the fuzzy partition matrix $U = (u_{ij})$ from the memory where a set of n training sample vectors $(x_1,\ldots,x_n)$ are stored. In this paper, j is represented as the vector index (j = 1, 2, …,n), where n is the number of training samples, the variable i represents the class index (i = 1, 2, …,C) , and C is the number of classes. For each training case x, its k nearest neighbours are identified by computing the Euclidean distances. The membership degree of the sample vector x_j in the class i is as given in Equation 1.

$$u_{ij}(x) = u_i(x_j) = \begin{cases} 0.51 + (n_i / k) \times 0.49, & if\ c(x_j) = i \\ (n_i / k) \times 0.49, & if\ c(x_j) \neq i \end{cases} \quad (1)$$

where n_i is the number of detected neighbours belonging to class i and $c(x_j)$ is the class label of the sample vector x_j . n_i denotes the labelled reference pattern among the k nearest labelled reference patterns labelled in class xi, and j ranges from 1 to n. Membership is distributed between classes. Based on this, it should be fairly obvious that u_{ij} is an element of a matrix U of size C x n. Equation 1 is used to assign higher fuzzy member values to training samples far from the decision boundary and lower fuzzy member values to patterns near the decision boundary (Keller *et al.*, 1985). Since u_{ij} is the fuzzy membership or sample value x_j of class i, u_{ij} must satisfy the following conditions:

$$U_{ij} \in [0, 1] \quad (2)$$

$$\sum_{i=1}^{C} u_{ij} = 1 \quad (3)$$

$$0 < \sum_{j=1}^{n} u_{ij} < n$$

Where

The FKNN algorithm then assigns fuzzy memberships of the unknown sample x to different classes according to the following equation:

$$u_i(x) = \frac{\sum_{j=1}^{k} u_{ij}(1/\|x - x_j\|^{2/(m-1)})}{\sum_{j=1}^{k}(1/\|x - x_j\|^{2/(m-1)})} \quad (4)$$

$$C^{(x)} = \arg \max_{i=1}^{C}(u_{i}(x_j)) \quad (5)$$

given that i = 1,2,…,C, and j = 1,2,..,k. In this context, j represents the j[th] sample vector among the k nearest neighbours of x. C is the number of classes; k denotes the neighbouring size. The fuzzy strength parameter m is used to ascertain how heavily the distance is weighted when calculating each neighbour's contribution to the membership value and is analytically determined from the training data. The value of m is usually taken as m ϵ (1, ∞). $\|x - x_j\|$ represents the distance which exists between x and its j^{th} nearest neighbour x_j. Where $u_i(x)$ is the membership of the test vector $\boldsymbol{x}$, to class i, $||\boldsymbol{x}-\boldsymbol{x}_j||$ is the Euclidean distance between the test vector $\boldsymbol{x}$, and the k-th nearest Neighbour vector $\boldsymbol{x}_j$, and m is a real number greater than or equal to 1.0 that sets the 'strength' of the fuzzy distance function. u_{ij}, which denotes the membership degree of the sample vector x_j in the class i, is computed in the first step of the algorithm.

In this study, features were extracted from the segmented training dataset and fed into the classifier where a user-defined class label is assigned to the obtained results. The FKNN classifier system possesses the capability to adjust itself to the ambiguity in the training and test data. The classifier ascribes a membership value to the unlabelled object and this provides the system with information which is useful in estimating the accuracy of the decision. This fuzzy membership value identifies what fraction of an unlabelled object resides in each of the defined classes. The FKNN classifier is implemented in a pair-wise manner for all the specified classes in the training set so that a reduction in the dimensionality of the feature-space can be achieved. The algorithm for the FKNN-based classification is given in Algorithm 1.

3.2 Locomotion Score Reference

An abridged version of the lameness scoring system proposed by Thomsen *et al.* (2008) was adopted in assigning threshold values for the standing, feeding, lying behaviours and the motion index. Each animal gets a score describing its behaviours with respect to lameness. The scoring system is presented in Table 1. The three different classes are 'lame', 'mildly lame' and

http://doi.org/10.46792/fuoyejet.v8i2.979 http://journal.engineering.fuoye.edu.ng/

'sound'. The observed difference is used in distinguishing between the behavioural scenarios.

Algorithm 1. The FKNN algorithm

Score	Term	Brief Description

Input: The dataset X for training with the labeled patterns $\{x_{ij}|_i = 1,2,...,n\}$ and the pattern y to test the classifier.
Output: Class label of y and confidence for each class label.
Step 1: For i = 1, 2,. . ., n
Step 2: Calculate the distance from x_i to y using the Euclidean distance;
Step 3: If x = i to k
Step 4: Include x_i in the set of kNNs;
Step 5: Else If (x_i is closer to y than any previous NNs)
Step 6: Dispose the farthest of the kNNs;
Step 7: Include x_i in the set of kNNs;
Step 8: End If
Step 9: End If
Step 10: End for
Step 11: For c = 1 to C
Step 12: Compute $u_i(x)$ using Equation (1);
Step 13: End For
Step 14: Assign a Crisp class label of y to the class with which has the highest membership value using Equation (5).

Table 1. A lameness scoring system
(An abbreviated version of Thomsen *et al*., 2008)

Score	Term	Brief Description
1	Sound	The animal walks normally, feeds properly. No signs of lameness.
2	Mildly lame	Cattle display some signs of lameness. Usually, an observer is not able to tell which limb is affected.
3	Lame	The animal is evidently lame on one or more legs. An observer can easily tell which of the legs is affected by lameness.

3.3 Description of the FKNN-based Model

The methodology is divided into three parts: data collection and preprocessing, feature extraction and classification. The proposed approach consists of two steps. We first performed stratified random sampling to reduce bias in the data, then normalized the data and performed feature reduction using PCA. After that, the FKNN-based model is first trained on the training set using stratified two-fold cross-validation (CV) to obtain the optimal pair of parameters (k, m) and then used to perform the resulting optimal FKNN model. It does classification work. The pseudocode for the PCA step is given by Algorithm 2.

Algorithm 2. Pseudocode for feature reduction procedure

Step 1: Load the data as a matrix X, whose rows represent instances, and columns connote features.
Step 2: Subtract the mean from each column belonging to X. The mean is the average across each dimension. This produces a new (normalized) matrix M whose mean is zero, such that it satisfies the required working conditions of PCA.
Step 3: Create a single value decomposition of M, namely, $M = U\sum V^T$, where U is an s*s orthonormal matrix consisting of the left singular vectors of M, $\sum$ is an s*t rectangular diagonal matrix with positive real numbers on the diagonal, and the t*t matrix V is the matrix of the right eigenvectors of M.
Step 4: Select a reduced dimension number L, project M down into the reduced space defined by only the first L singular vectors U_L, then the new matrix $N = (U_L)^T M$ is obtained.

3.4 Classification Phase using FKNN

In the classification phase, the model does its classification using the new feature set produced by the PCA. The first step is to set up all the model parameters. An experimental outline was drawn to select the optimal fuzzy strength parameter for the FKNN classifier. An interval of (1, 2) was chosen and continually increased with steps of 0.1 for the fuzzy strength parameter m, and 2-fold CV was performed using various values of nearest neighbours k. For values of m in the interval (1, 2), the average accuracy obtained by FKNN by the way of Cross Validation, and eventually, the one with the highest average accuracy was selected as the optimal fuzzy strength parameter. After which the FKNN classifier was employed in computing the classification accuracy using the reduced feature set. The pseudo-code for the classification phase is presented in Algorithm 3.

4 Results and Discussion

The model simulation was implemented using Matrix Laboratory (MATLAB) 2016. The proposed model uses 285 data cases for model training and 285 data cases for testing. The dataset is a synthetic dataset comprised of a range of physiological measurements of the lying, standing, feeding behaviours and overall motion index from of a total of 570 cows, 280 sound, 189 mildly lame and 101 lame. Normalization was carried out on the data in order to stop the feature values in greater numerical ranges from dominating those in smaller numerical ranges. These patterns are grouped into three main categories depending on their membership strengths. The system assigns a user-defined label (in the range of 1 to 3) for each pattern analysed. The accuracy obtained from a two-fold cross validation is presented in Table 2. The resulting models for three separate optimal test cases (when m=1,1.88 and 2.00), the resulting models are tagged FKNN-D, FKNN-E, FKNN-F respectively as presented in Table 3.

http://doi.org/10.46792/fuoyejet.v8i2.979 http://journal.engineering.fuoye.edu.ng/

Algorithm 3. Pseudocode for Classification Phase

```
performance estimation by using n-fold CV
where n = 2*/
Begin
    For i = 1 to Mmax
        For j = 1 to k
          Training set = k - 1 subsets;
          Test set = remaining subset;
          Train the FKNN model on the
            training set to find the optimal fuzzy
           strength parameter m when the
             neighbourhood size k is set to 3, 5,
           7,8and 9 respectively;
           Test it on the test set and assigns the
           accuracy to V(j),
               where V is a vector whose element
           is the corresponding accuracy
           obtained by each folder;
        End for
        Compute the mean value of vector V,
         and store the mean CV accuracy in the
        vector M(i);
    End for
    Find the optimal m value whose
      corresponding mean CV accuracy is the
     highest in M(i);
End
Begin
    For l = 1 to k
        Training set = k - 1 subset;
        Test set = remaining subset;
        Train the FKNN model on the training
        set using the obtained optimal
        parameter combination;
        Test it on the test set and save the mean
        CV accuracy;
End for
Return the average classification accuracy
  rates of FKNN over l test set.
End
```

When the number of features were 4, the obtained correct classification rates were 89.56%, 89.74% and 89.21% respectively and in the second test, when the features were 3, for the selected k, maximum values of 94.03%, 93.33% and 92.97% were obtained for the three different models, as presented in Table 2. The classifier generates tables for membership values for all values of KNNs prior to segmentation and classification so that elusive lameness patterns can be detected and correctly classified. For each model, a uniform accuracy value was obtained across both k-fold iterations of the test. The classification accuracy fluctuates between 92.97% and 94.03% for a feature set of 4 with different values of m. The descriptive statistics of models FKNN-D, FKNN-E and FKNN-F are presented in Table 3.

It can be seen from the table that the fuzzy strength parameter has a great impact on the performance of the FKNN classifier. For the first set of experiment, the best classification accuracy was obtained when m was 1.00, 1.88 and 2.00 respectively on feature space of 3, as presented in Table 3.

Table 2. Accuracies obtained by the FKNN classifier in Stratified 2-fold CV

FKNN	Performance Metric
FKNN-D	Accuracy (%) 94.03
k=4.00	Sensitivity(%) 96.74
m= 1.00	Specificity(%) 93.74
	AUC (%) 95.00
FKNN-E	Accuracy (%) 93.33
k=5.00	Sensitivity (%) 91.43
m=1.88	Specificity (%) 92.16
	AUC (%) 94.86
FKNN-F	Accuracy (%) 92.97
k=3.00	Sensitivity (%) 90.72
m=2.00	Specificity (%) 90.62
	AUC (%) 94.50

Table 3. Descriptive Statistics of Models FKNN-D, FKNN-E and FKNN-F

Set	Features	Fold 1	Fold 2	Average Accuracy
Training Set	3			
Test set 1	3	94.73%	93.33%	94.03%
Test set 2	3	94.03%	92.63%	93.33%
Test set 3	3	94.38%	91.57%	92.97%

The results confirm that a reduced feature space yields higher accuracy than the original feature space, as seen in Table 3, the accuracy fluctuates between 91.57 and 94.73 for both tests.

5 Conclusion and Future Work

This work will be useful in reducing economic losses for cattle producers by proposing a model which will help improve animal healthcare and farm profitability. There's a need for a larger varying data obtained across a number of farms or animal groups as this can improve the early detection of lame cattle in the future.

References

Afonso, J.S., Bruce,M., Keating,P. Raboisson, .,Clough, H., Oikonomou, G. and Rushton, J. (2020). Frontiers. Profiling detection and Classification of Lameness Methods in *British Dairy Cattle Research:* A Systematic Review and Meta-Analysis.

Alsaaod, M.; Fadul, M.; and Steiner, A.(2019). Automatic lameness detection in cattle. *Vet. J.* 2019, 246, 35–44.

Chapinal, N.; De Passillé, A.; Rushen, J.;and Wagner, S.(2010). Automated methods for detecting lameness and measuring analgesia in dairy cattle. *J. Dairy Sci.* 2010, 3, 2007–2013.

Duda, R. O., and Hart, P. E., (1973). Pattern classification and scene analysis. Wiley and Sons, New York. 184-189.

Gardenier, J.; Underwood, J.; Weary, D.; and Clark, C. (2021). Pairwise comparison locomotion scoring for dairy cattle. *J. Dairy Sci*. 2021,104, 6185–6193.

Herrero, M.; Henderson, B.; Havlík, P.; Thornton, P.K.; Conant, R.T.; Smith, P.; Wirsenius, S.; Hristov, A.N.; Gerber, P.; and Gill, M. (2016). Greenhouse gas mitigation potentials in the livestock sector. *Nat. Clim. Chang*. 2016, 6, 452 461.

Keller, J. M., Gray, M. R., and Givens, J. A., (1985). A fuzzy k-nearest neighbours algorithm. *IEEE Transactions on Systems, Man, Cybernetics*, Vol.15, 580–585.

Liu, J.; Dyer, R.M.; Neerchal, N.K.; Tasch, U.; and Rajkondawar, P.G. (2011). Diversity in the magnitude of hind limb unloading occurs with similar forms of lameness in dairy cows. *J. Dairy Res*. 2011, 78, 168–177.

O'Leary, N.; Byrne, D.; O'Connor, A.; and Shalloo, L. (2020). Invited review: Cattle lameness detection with accelerometers. *J.Dairy Sci*.2020, 103, 3895–3911.

Pluk, A.; Bahr, C., Poursaberi, A , Maertens, W., Van Nuffel, A. and Berckmans, D. (2012). Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. *Journal of dairy science. 95. 1738-48.* 10.3168/jds.2011-4547.

Schlageter Tello, A.,Van Hertem, T.,Viazzi, S., Groot, K., P.W.G., Bahr, C., and Lokhorst, Kees. (2015). Relation between observed locomotion traits and locomotion score in dairy cows. *Journal of Dairy Science.* 10.3168/jds.2014-9059.

Sjöström, K., Fall, N., Blanco-Penedo, I., Duval, J.E., Krieger, M. and Emanuelson, U. (2018). Lameness prevalence and risk factors in organic dairy herds in four European countries. *Livest. Sci., 208 (2018)*, 44-50 https://doi.org/10.1016/j.livsci.2017.12. 009

Thomsen, P. T., Munksgaard, L., and Tøgersen, F. A., (2008). Evaluation of a lameness scoring system for dairy cows. *Journal of Dairy Science*, Vol. 91(1), 119–126.

Thorup, V.M.; Munksgaard, L.; Robert, P.E.; Erhard, H.; Thomsen, P.; and Friggens, N.(2015). Lameness detection via leg- mounted accelerometers on dairy cows on four commercial farms. *Animal* 2015, 9, 1704–1712.

Van Hertem, T., Viazzi, S., Steensels, M., Maltz, E., Antler, A., Alchanatis, V., Schlageter-Tello, A.A., Lokhorst, K., Romanini, E.C.B., Bahr C., Berckmans, D., and Halachmi, I., (2014). Automatic lameness detection based on consecutive 3D-video recordings. *Journal of Biosystems Engineering,* Vol.119, 108 116.

Viazzi, S.; Bahr, C.; Van Hertem, T.; Schlageter-Tello, A.; Romanini, C.; Halachmi, I.; Lokhorst, C.; and Berckmans, D.(2014).Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Comput. Electron. Agric*. 2014, 100, 139–147.

http://doi.org/10.46792/fuoyejet.v8i2.979 http://journal.engineering.fuoye.edu.ng/