Image-based Individual Cow Recognition using Body Patterns

Rotimi-Williams Bello¹, Abdullah Zawawi Talib², Ahmad Sufril Azlan Mohamed³
Daniel A. Olubummo⁴, Firstman Noah Otobo⁵

School of Computer Sciences, Universiti Sains Malaysia, 11800, Pulau, Pinang, Malaysia¹,²,³
Department of Computer & Information Systems, Robert Morris University, Moon-Township, Pennsylvania, USA⁴
Department of Mathematical Sciences, University of Africa, Toru-Orua, Bayelsa State, Nigeria¹,⁵

Abstract—The existence of illumination variation, non-rigid object, occlusion, non-linear motion, and real-time implementation requirement has made tracking in computer vision a challenging task. In order to recognize individual cow and to mitigate all the challenging tasks, an image processing system is proposed using the body pattern images of the cow. This system accepts an input image, performs processing operation on the image, and output results in form of classification under certain categories. Technically, convolutional neural network is modeled for the training and testing of each pattern image of 1000 acquired images of 10 species of cow which will pass it through a series of convolution layers with filters, pooling, fully connected layers and softmax function for the pattern images classification with probabilistic values between 0 and 1. The performance evaluation of the proposed system for both training and testing data was carried out for each cow’s identification and 92.59% and 89.95% accuracies were achieved respectively.

Keywords—Cow; body patterns; convolutional neural network; image; recognition

I. INTRODUCTION

Cows in the past were classically monitored with the sole aim of aiding tracking, health information, performance recording, prevention against manipulation and swapping, and verification of false insurance claims. There are basically two recognition techniques employed for the identification of the animal. One recognition technique leaves a permanent mark on the animal for identification while the other recognition technique leaves a temporary mark. Examples of the recognition technique that leaves a permanent mark are found in [1], [2], [3], [4], [5] with their drawbacks. The tattooing of ears, tagging of ears, microchips implant and branding are popular invasive identification techniques that leave a permanent mark on the animal’s body with so many challenges such as animal infections, mild sepsis, and hemorrhaging [2], [3].

Examples of the recognition technique that leaves a temporary mark on the body of the animal for identification purposes are found in the work of Barron et al. [6] with their drawbacks. Among the classical methods of animal identification are drawing, tagging, tattooing, branding, notching, and Radio Frequency Identification (RFID). However, classical methods of animal identification have notable adoption problems which have contributed to the low acceptance rate of the methods among the cow breeders. The classical methods of animal identification are not reliable; they are prone to fraudulent activities such as swapping, duplication and forgery of the so called unique identification numbers tagged on the animal’s body [7], [8], and therefore cannot meet the required level expected from them for the monitoring and identification of animal [9].

Many automatic systems have been proposed recently for the monitoring and identification of cow however, most of these devices are sensor based and sometimes become burden and injurious when worn on the body of the animal [10]. There is need for automatic cow monitoring system in livestock farm to be developed as there is uprising in the number of cow year in year out in almost every part of the world and there is great task involved in monitoring cow manually. Lu et al. [11] proposed cow traceability system that was based on the iris analysis for the enhancement of cow management. The image quality assessment of the captured iris sequences was firstly made before the clear iris was selected. By using segmentation that was based on edge detection, the inner and outer boundaries of the iris of the cow were fitted as ellipse form. The iris image was normalized using geometric method and both the local and the global features of the iris of the cow were extracted using 2D complex wavelet transform. However, in an unconstrained environment where there is greater possibility of getting poor quality image of cow’s iris, this method may not be appropriate for a reliable traceability.

By using video data, there is every possibility that the problems attributed to the classical methods can be mitigated using the visual based automatic cow recognizing system. The recognition of individual cow in the automatic cow monitoring implementation process enables behavior monitoring of individual cow at long run for body condition score which plays important role in the health condition of individual cow. The system proposed in this paper is image-based individual cow recognition using body patterns. The rest of the paper is organized as follow. Presented in Section 2 are the literature review, followed by the material and method in Section 3, the results and discussion are in Section 4. The conclusion is in Section 5.

II. LITERATURE REVIEW

The conventional constructs of identifying animal can be categorized into: (1) permanent recognition construct (PRC);
(2) semi-permanent recognition construct (SRC); and
(3) temporary recognition construct (TRC); [12, 13]. The
tattooing of ear and body, tagging of ear, microchip implants
and branding are referred to as PRC recognition methods [14]
but with several limitations [15] such as: (1) lack of large
scale production of various metal clips and plastic tags that
can be enough for the identification of large-scale animal;
(2) easy lost of the available ear tags due to ear tearing;
(3) infections of animals such as cattle and other ruminant
animals due to notches [16], [17], [18], moreover, more than
half percentage of the animals are infected from the injury
sustained on their ear due to the implanted plastic ear-tags,
reason being that, the ear-tags cause various health challenges
such as local inflammation, thickening of the flesh, presence
of pus-forming bacteria, and loss of blood through the notch
[17], [13]. Cattle recognition using methods such as pattern
sketching and collar is SRC method. Furthermore, the use
of dye or paint and radio frequency identification (RFID) based
recognition are referred to as TRC for the recognition of
animal [12], [19].

According to [20], the sketching pattern is applied for
the recognition of cattle such as Holsteins and Guernsey with
broken color. High drawing skills of an individual for
sketching is needed which should be comparable to standard
image quality and positively affect the cattle identification
process. However, this method cannot be used for the
identification of solid collared breeds such as Red Poll and
Brown Swiss breed as some artificial marking methods such
as ear tagging and tattooing that are discrimination based are
needed. However, the method of ear tagging damages the
cattle’s ear at the long run. As iterated in Petersen’s work [21],
muzzle print-based cattle recognition method using blue ink
and A-5 paper [22] was the first attempt to get permanent
recognition method for cattle. In the method, skills are
required to acquire the muzzle pattern’s print image, by
holding firm the cattle.

Lately, the research community has shifted attention to
advancing cattle recognition using image of muzzle print as a
new paradigm for cattle identification [22], [20]. According to
[23], print image of muzzle pattern is made up of beads and
ridges patterns. Muzzle dermatoglyphics such as granola,
ridges, and vibriasa from various breeds are not the same
[16]. Similarly, proposed in Mishra et al. [24] is method of
cattle breeds recognition using the beads and ridges features of
muzzle print images. Similar to the work of Mishra et al. [14]
is Minagawa et al. [22], they proposed a cattle identification
method using muzzle print, the performance evaluation was
made using filtering techniques for muzzle image analysis and
morphological approaches. Equal Error Rate (EER) of 0.419
was reported by them.

Contrary to Minagawa et al. [22] is a framework proposed
by Barry et al. [25]. The framework is a cattle recognition
using muzzle print images. They reported the 241 false non-
match rates (FNMR) over 560 genuine acceptance rate (GAR)
and 5197 false matches over 12,160 impostors matching
closely with the same value of EER of 0.429, respectively.
In their cattle identification effort, Kim et al. [26] proposed a
method that could recognize the Japanese black cattle using
the cattle face’s pixel intensity [26]. Proposed in [27] is a local
binary pattern (LBP) based model for recognition of cattle
using the texture features of cattle facial representation.
Proposed in [28] is an approach for cattle recognition based on
Speeded Up Robust Feature (SURF) descriptor. The approach
was an enhancement of Petersen’s method for cattle
identification. The results of experiment was reported based
on the image datasets of 4 cattle breeds which were used to
captured on A-5 paper with blue inked for the purpose of
cattle recognition. Proposed in [20] is a matching refinement
technique in scale invariant feature transform (SIFT)
descriptor for cattle recognition using database of 160 muzzle
print images. By the application of matching refinement
technique in SIFT approach, the matching scores of the key-
points of muzzle print images were computed. Nevertheless,
the performance of the matching refinement approach and the
original SIFT approach were compared, and the value of EER
equal to 0.0167 was achieved.

Proposed in Awad et al. [29] is a framework for
recognizing cattle using SIFT descriptor approach. The
approach is used for localizing and detecting the beads and
ridges’ key points in the images of muzzle print for the cattle
identification. The RANdom Sample Consensus (RANSAC)
technique incorporated in the SIFT algorithm is used for the
palliation of the outliers in muzzle image for an improved,
robust, and reliable cattle identification. Database of 90
muzzle images was used for the experiment where 15 muzzle
images were captured from each cattle of 6 in number.
Proposed in Tharwat et al. [23] is an approach of cattle
recognition that was based on muzzle image using the
technique of local texture descriptor. The technique works in
such a way that the texture extraction algorithm that was based
on local binary pattern used the local texture features
extraction from the images of muzzle point. The involvement
of more processing time in the cattle recognition process is a
major limitation of the technique.

Object recognition method that is based on CNN was
proposed in [30]. The proposed architecture which combines
RGB image and its corresponding depth image for object
recognition is made up of two unconnected CNN processing
streams, which are sequentially integrated with a late fusion
network. ImageNet [31] is employed for the training of the
CNNs in which the depth image is encoded as a rendered
RGB image, making the information that is contained in the
deep data to go round over all the three RGB channels, and
subsequently, a standard and pre-trained CNN is employed
for the recognition. Due to limited availability of large scale depth
datasets that are labeled, CNNs that are pre-trained on
ImageNet [32] are employed. Proposed in [33], is another
object recognition method, which employs deep CNN. The
proposed method also uses CNN, which is pre-trained for
image classification and provides a robust, semantically
meaningful feature set. The depth information is integrated by
rendering objects from a canonical perspective and getting the
depth channel colorized according to distance from the object
center.

Jingqui et al. [34] proposed the method of object
recognition based on image entropy; this was aimed at
identifying the behavior of cow object that is on the motion
against a complicated background. They used the minimum
bounding box and contour mapping for the real-time capture of behavioral and characteristic features displayed by the cow. Although the approach used has time-saving advantage for cow breeders and yields a high recognition rate of estrous and hoof-disease not less than 80%, the time correlation of cow behaviors was not integrated.

Andrew et al. [35] demonstrated the suitability of computer vision pipelines that utilize deep neural architectures to carry out automated Holstein Friesian cattle detection in addition to individual identification in a farm set up. They showed that it is possible to perform robustly Friesian cattle detection and localization with an accuracy of 99.3% on the available dataset. Although they showed the capability of their method in the scenarios presented, they did not consider complicated setups such as faster moving, larger herds and tight animal gatherings.

In the process of extracting features from an image, Kumar et al. [36] posited that pre-processing is important for object tracking accuracy but feature extraction and representation algorithms that are based on appearance are unable to perform the recognition of object as a result of image blurriness due to noise, low illumination and the unconstrained environment under which the images were captured. Therefore a method based on feature descriptor techniques is utilized for the unique identification of individual object. Based on the pre-processing process, reliable results were obtained from the tracking process of the object. Pre-processing which majorly involves particle filtering and segmentation of muzzle point images is necessary in the features extraction process. The primary aim of undergoing pre-processing of the muzzle images using enhancement algorithms before the feature extraction and matching process take place is to ensure that the muzzle images are enhanced before the analysis of the extracted texture features and for better representation in the feature space.

III. MATERIAL AND METHOD

A. Equipment for Experiment

Ten (10) species of cow were examined in recognizing the characteristic of individual cow, each having 100 images making 1000 images in total. The patterns of the black and white body of the 10 species of cow were used for the calculation of the input parameters values for training. 400 images of body patterns (40 cows (subject) × 10 images of each subject) were used for the training of the proposed deep learning approach in the training phase. 600 pairs of testing (60 cows (subject) × 10 images of each subject) of the body patterns images in each fold were used for testing the probe images in the testing phase. By middle of September 2018, a test was performed in order to get the image data and the image data was analyzed accordingly by image processing. A charged coupled device (CCD) camera was employed for the side image capture of each cow. In order to obtain images of required width (235-270cm), the CCD camera was placed on a high pole away from the experimental system centerline. The image processing system was strategically placed in a location through which the cow passed everyday with minimized illumination variation for the production of noiseless and clear images as shown in Fig. 1.

The cow recognition and identification system can run on any Windows-based personal computer. A faster computer system is recommended for the processing of the images that involves calculations and processing on the go. The personal computer specifications for development of the cow recognition and identification system are Intel core i5 Processor, 8 Gigabyte of RAM, Graphics card, 2 terabytes of hard disk space, a CCD digital camera, and a computer monitor for digitizing, displaying, and processing multiple images. The specification for the execution of the image-processing and computer vision elements is OpenCV and its library.

B. Processing of Images

The filtration technique used for this work is Gaussian filtering technique while multi-layer deep learning neural network was used as a classifier for the cow identification and contrast limited adaptive histogram equalization (CLAHE) was used for enhancement of the contrast between the cow’s body patterns. The difference of the Gaussian filter was got by finding the difference between two Gaussian functions [37].

Fig. 2 shows some image samples of cow’s body patterns from the database. Fig. 3 shows the database containing blurred image patterns of the cow’s body affected by the unconstrained environment and postures of the cow leading to poor quality of the images. Using Norouzzadeh et al. [38], we filtered the images to get rid of the blurriness, background patches and low illumination.

Fig. 1. Dimensional Sketch of the Individual Cattle Recognizing System.

Fig. 2. Images of Cow’s Body Patterns from the Database.
In order to enhance the identification process and remove the patches and the noises from the captured images that were collected, various image pre-processing techniques were applied. Low illumination and poor image quality are the most two fundamental challenges confronting image acquisition especially images of cow’s body patterns. The images captured in an unconstrained environment were converted to grayscale images in order to reduce the patches and the noises captured with them. The converted images were improved upon by contrast limited adaptive histogram equalization based image processing technique.

The pre-processing technique accepts the images in their color form and converts them to grayscale before fetching them into the filter for removal of the patches and the noises contained in the captured images. The feature extraction involves the convolution and pooling operations on the images until the images get to the classifier for classification analysis for the generation of the desired output (Fig. 4). The removal of the noises was carried out using an auto-encoding technique. Stacked denoising auto-encoder (SDAE) technique initializes deep network and it is applicable for encoding and decoding the texture features of the image patterns that were extracted and encoding the extracted sets of features for optimum representation of the feature [39].

Technically, convolutional neural network (CNN) is modeled for the training and testing of each input image which will pass it through a series of convolution layers with filters, pooling, fully connected layers and softmax function for the image classification with probabilistic values between 0 and 1. As shown in Fig. 4, the first layer to extract features from the input image is convolution. Convolution primarily conserves the relationship between pixels by learning the image features using squares of input data. It involves a mathematical operation with two inputs such as image matrix and a filter. When there are too large images, pooling layers primarily reduce the number of parameters (dimensionality size). In the proposed CNN as seen in Fig. 4, the operation of the pooling is applied individually to each feature map.

Generally, the more the convolutional steps become, the more the complex features possibility of being recognized becomes using the proposed network. Until the system can dependably recognize objects, the whole process is repeated in successive layers. Each layer’s neurons of the CNN as seen in Fig. 4 are in 3D arrangement, making a transformation of a 3D input to a 3D output. For instance, for an input image, the first layer which is the input layer takes the images as 3D inputs, with width, height and color channels as the dimensions of the image. The first convolutional layer’s neurons connect to the input images’ regions and change them into a 3D output. Each layer hidden units learn nonlinear combinations of the original inputs which becomes the inputs for the layer that follows. By this, at the end of the network, the learned features become the inputs to the classifier.

The intensity values of the gray scale of the background images are more than 100 but less than 150 in respect to the colors of the cows’ body surface. 128 was fixed as the pixel’s threshold value for the whole image. While 1 is assigned as the binary values for the intensities that are greater than the threshold value of 128, 0 is assigned as the binary values for the intensities that are less than the threshold value of 128. Because the threshold value could be changed with illumination and noise, it becomes very important. Individual cow’s image is captured for the identification of their individual characteristics. Individual cow identification using unique body patterns is made possible because of the invariant of the body patterns to growth. This uniqueness enables the patterns to be used as the input layer values in the neural network algorithm.

IV. RESULTS AND DISCUSSION

Having tried out the effectuality of the proposed approach using images of cow’s body patterns for the recognition and identification of the cow, the comparison with other recognition algorithms is attained in order to evaluate the accuracy of the identification in proliferation settings. Evaluating the performance of the experimental results, the database of the cow’s body images is segmented as follows: (1) the training phase; and (2) the testing phase. 400 body images of different subjects (40 cows (subject) × 10 images of each subject) were used for the training of the proposed approach in the training phase. 600 pairs of testing (60 cows (subject) × 10 images of each subject) of the body patterns images in each fold were used for testing the probe images in the testing phase.
The second part is “pre-training.” For a given training set of image data \( \{x_1, x_2, ..., x_N\} \), the learning system obtains a model through learning (or training) to describe the mapping relationship between input and output variables. This work assumed that RBM model has this descriptive ability, therefore it consists of several layers, through which the input is the image expression data vector while the output is the abstracted higher-level feature vector. Each layer of RBM networks undergoes individually unsupervised training to ensure that feature information is preserved to the uttermost as feature vectors are mapped to different feature spaces. To construct the joint distribution model of visible layer and the hidden layer through energy function, the joint probability maximum likelihood of training sample under model parameter \( \theta \) is calculated by

\[
P(v|\theta) = \frac{1}{Z(\theta)} \sum_h e^{-E(v,h|\theta)}
\]  

(3)

The third part is “fine tuning.” Fine-tuning is a common strategy in deep learning to carry out supervised learning through tagged sample training set \( T = \{(x'_1, y'_1), (x'_2, y'_2), ..., (x'_N, y'_N)\} \). After that, the top feature vectors corresponding to sample output by the multi RBM network are formed based on the training set of statistical classification structure. This part is a BP network; it takes a specific dimension feature vector to a softmax function. In order to get the best connection weights, this work considered solving the following optimization problem using particle swarm optimization (PSO), so that the loss of function in the training set is minimized.

\[
\hat{\theta} = \arg \max_{\theta} \frac{1}{N} \sum_{n} l(\theta, x_n, t_n)
\]  

(4)

The last part is the “class identification.” Tested sample \( x_{N+1} \) as network input is subjected to feature learning and abstraction through a network model training to produce a corresponding output \( y_{N+1} \) by

\[
y_{N+1} = \arg \max_p (y_{N+1}|x_{N+1})
\]  

(5)

and thus achieve classification.

For the evaluation of performance, the local feature descriptor technique was used to extract and encode texture features of the cow’s body patterns. As earlier mentioned, the normalization and the descriptor process help in mitigating the external factors such as low illumination, poor image quality, and background patches affecting the captured images. In performing the tasks involved in this process, cells are converted to blocks. During this process, blocks are overlapped and cells shared among the blocks and normalized separately. Scale-invariant Feature Transform (SIFT) and Rectangular-Histogram of Oriented Gradients (R-HOG) are similar though, they don't align to their dominant orientation (Fig. 8(b)). SDAE produced the best experimental results (Fig. 6 and Fig. 7) when compared to other approaches used in this work making it fit the most for the denoising. 400 body images equivalent to (40 cows (subject) \times 10 \text{ images of each subject}) were chosen randomly for system training and 600 body images equivalent to (60 cows (subject) \times 10 \text{ images of each subject}) were used for the testing. The experimental results are reported and analyzed as found in Table 1.

For the training of the proposed deep learning framework using deep belief network (DBN) as shown in Fig. 5, there is a need for a monolithic database amount. Although the number of cow’s body images in the database is encouraging, it is not satisfactory enough to train the stacked denoising auto-encoder with a database of 1000 worth of cow’s body patterns images. Therefore a transfer learning approach is needed to fine-tune the weight between the input and the hidden layer and determine the pre-training of the proposed deep learning approach.

The basic mathematical steps that are involved in using the deep belief network for this work are as follows:

- **Problems setting:** Given a training set of pre-processed body pattern image data
  \[ T = \{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\} \]  
  \( (x_i, y_i), i = 1, 2, ..., N \) denotes the sample point, \( x_i \in X \subseteq \mathbb{R}^n \) is the sample image data whereas \( y_i \in Y \) is the corresponding tag of the label; the recognition procedure of proposed system is to input data set \( T \) to the network, find the mapping between input \( X \) and output \( Y \) to form a generative joint probability distribution model formula \( P(X, Y) \), generate the output \( y_{N+1} \) by
  \[
y_{N+1} = \arg \max_y p(y_{N+1}|x_{N+1})
\]  
  for a given prediction sample \( x_{N+1} \), and judge the image classification of \( x_{N+1} \) according to \( y_{N+1} \). The system contains the following parts as shown in Fig. 4:

- The proposed cow’s body patterns image identification using deep belief network and a back propagation (BP) network layer, wherein the multi-layer RBM is used to input data feature learning to achieve abstraction and dimensionality reduction of data through the hierarchical feature learning as is shown in Fig. 5; BP network layer is a categorical network, and it is to categorize the abstracted higher-level features through softmax function. The softmax function, also known as softargmax or normalized exponential function, is a function that takes as input a vector of \( K \) real numbers, and normalizes it into a probability distribution consisting of \( K \) probabilities proportional to the exponentials of the input numbers.

The first part of the processes as shown in Fig. 4 is “preprocessed cow’s body patterns images” which are introduced as inputs to the proposed networks for features extraction and classification.

Fig. 5. RBM-based DBN Model.
The significant reason for binary patterns (Fig. 8(a)) is to sum up the local structure in a block through comparison of each pixel with its neighborhood [40]. Each pixel coded with a sequence of bits is colligated with the connection between the pixel and one of its neighbors. The center pixel’s intensity is denoted with 1 if it is greater than or equal to its neighbor, and denoted with 0 if otherwise with a binary number at the end created for each pixel.

V. Conclusion

Image-based individual cow recognition using body patterns was the main work carried out in this research. Cows usually are identified to prevent them from being stolen or protect them from danger, and in many agricultural settings, their behaviors are usually studied using imaging technology to enable timely monitoring and identification of health challenges. CNN and some other popular image recognizing techniques such as DBN, SDAE, CLAHE, Gaussian filter, binary pattern, were employed in this work for the cow recognition. The various techniques were discussed in details as they are applicable to the cow recognition process. Datasets of 1000 images of cow’s body patterns from 10 species of cow were created for this work where 400 images were employed for the training and 600 images were used for the testing. The advantage of using this datasets is the various species of cow whose images are contained in the database used for the recognition. Gaussian filtering technique was used as the filtration technique; this was supported by SDAE for denoising while multi-layer convolutional neural network was used as a classifier in comparison to deep belief network which needs a monolithic database amount for the cow identification, and contrast limited adaptive histogram equalization (CLAHE) was used for enhancement of the contrast between the cow’s body patterns. The performance evaluation of the proposed system for both training and testing data was carried out for each cow’s identification and 92.59% and 89.95% accuracies were achieved respectively. Although this work has been able to apply modern image-based identification method for the recognition of cow using body patterns, recognition of occluded and non-linear moving object such as cow in real-time using the object’s multi-features is a work that we consider worthy of investigating in the future.

ACKNOWLEDGMENT

This work was supported in part by the Institute of Postgraduate Studies, Universiti Sains Malaysia, 11800 USM, Penang, MALAYSIA.

REFERENCES


