



Segmentation and detection of cattle branding images using CNN and SVM classification

Juliano Weber^a, Bruno Belloni^b and Carlos Silva^c

^a Federal Institute Farroupilha, Santo Angelo, Brazil

^b Federal Institute Sul-Rio-Grandense, Passo Fundo, Brazil

^c Federal Institute Farroupilha, Alegrete, Brazil

juliano.weber@iffarroupilha.edu.br, bruno.belloni@hotmail.com, carlos.santos@iffarroupilha.edu.br

KEYWORD

Convolutional Neural Network; Support Vector Machines; Cattle Branding

ABSTRACT

This article presents a hybrid method that uses Convolutional Neural Networks (CNN) to segmentation and Support Vector Machines (SVM) to detection the brandings. The experiments were performed using a cattle branding images. Metrics of Overall Accuracy, Recall, Precision, Kappa Coefficient, and Processing Time were used in order to assess the proposed tool. The results obtained here were satisfactory, reaching a Overall Accuracy of 93% in the first experiment with 39 brandings and 1,950 sample images, and 95% of accuracy in the second experiment, with the same 39 brandings, but with 2,730 sample images. The processing time attained in the experiments was 32s and 42s, respectively.

1. Introduction

The use of computer tools to aid in the analysis and recognition of images is a subject of interest for the most renowned research centers in the world. Its use for image analysis and recognition is in constant development, generating many benefits to society in the most diverse areas of knowledge. Concerning the recognition of cattle branding in particular, a very traditional activity that is highly economically and socially relevant for Latin American countries, including Brazil, there is no specific and appropriately consolidated tool for this end. As an example of the importance of this activity, according to the Food and Agriculture Organization – FAO, among the producing countries, Brazil and India have the largest herds, Brazil being the 1st, with an average of 209,215,666 cattle heads (Atlas, 2015).

The use of brands or symbols on cattle presupposes the public recognition of its property by an individual or group. Livestock has a relevant role in the shaping of society. To these days, it is still an activity of great importance for cultural expressions, as it is linked to the culture and the countryside way of life. It also plays a role in the affirmation and building of individual and group identities (Arnoni *et al.*, 2013). The use of cattle branding dates back to the beginning of the Iberian colonization in America.

The beginning of its institutionalization happened when they started being registered in official agencies, the recognized holders of public legitimacy (Arnoni *et al.*, 2013). These records followed regulations that seek to legitimize branding, as well as regulate the manner and timing of the procedure, discriminate how the records are



made, assign fees to the records, regulate the craft of the irons, and government taxes. In Brazil, attempts and investments to upgrade the cattle branding recording system were always subject to controversy, due to opposition from agriculturists. A major part of their concern is associated to a fear of losing family brands and the meaning they have acquired through time. Currently, brand recording in Brazil is performed by town offices, generally without a more effective systematization and without instituted renewals. Generally, cattle brand records involve books with drawings of the brands and the identification of their owner.

In Brazil, attempts and investments to upgrade the cattle branding recording system were always subject to controversy, due to opposition from agriculturists. A major part of their concern is associated to a fear of losing family brands and the meaning they have acquired through time. Currently, brand recording in Brazil is performed by town offices, generally without a more effective systematization and without instituted renewals.

In face of the presented scenario, this work intends to present and assess a tool that performs automatic cattle brand recognition, with the goal of replacing the manual control of cattle branding performed today. Our objective is to potentially decrease the possibility of duplicate records, reducing waiting times for the recording of new brands, improving governmental administration regarding the brand archive under its care, and aiding security officers in the prevention of cattle raiding crimes.

Our work focused on the development of algorithms for the recognition of cattle branding for the City Hall, in Rio Grande do Sul, Brazil. The text is organized as follows: in Section 2, we describe related works. In Section 3, we present the materials and methods used in our approach. Section 4 describes the results and discussions obtained through the application of the recommended methods. Lastly, in Section 5, the final considerations are presented.

2. Related Work

In general, we could not find any works in the literature review that report the use of convolutional neural networks for the recognition of cattle branding images. Sanchez *et al.* (2016) present a tool for recognition of cattle branding that uses Hu and Legendre moments for extracting features of images in a gray scale, and also a classifier of k-nearest neighbors (k-NN). The authors used a set containing 40 brandings and a total of 100 images. The method proposed by the authors presented a significant loss of accuracy when applied to a large number of images. Another disadvantage was the need for a pre-processing of the original branding images used in the experiments, in order to correct noise and imperfection issues.

Silva *et al.* (2017) present a tool for recognition of cattle branding that uses CNN and SVM. The authors used a cattle branding image set offered by the City Hall of São Francisco de Assis, Brazil. In the experiments, the authors used 12 cattle brandings, and a set of 540 images. The accuracy obtained in the experiment reached 93.28%, but the authors highlighted the need to increase the number of sample images in order to validate the proposed method's ability to correctly classify large volumes of brandings, without compromising the accuracy or dramatically increasing the computational cost of identifying cattle brandings.

Differently from the research presented by Silva *et al.* (2017), the present study shows new experiments and results using the hybrid method combining CNN and SVM on a new set of cattle brandings. Moreover, the images were provided in a smaller format (253x253 pixels) and colored using the RGB color model. We also performed a comparative study between both experiments, using different quantities of samples for 39 brands, in order to establish a correlation between the increase in the overall accuracy of the method and the number of sample images.

Hybrid methods are being proposed and exposed in the literature for solving problems that involve feature extraction and classification of digital images, as exemplified by the study by Niu *et al.* (Niu *et al.*, 2011), where the authors present an algorithm using convolutional neural networks and support vector machines to solve the issue of text recognition. The authors assessed the model from two different perspectives: recognition performance and reliability. The results obtained in the experiments were better than those obtained with the application of traditional methods, thus, the authors concluded that a combination between CNN and SVM yielded more satisfactory results, with a recognition rate of 99.81% without rejection, and a recognition rate of 94.40%, with 5.60% of rejection.

The works found in the literature about recognition and classification of images using descriptors or filters for extracting features, followed by a step dedicated to quantization and grouping, and, finally, a classification stage are divided in two categories: algorithms with a single stage for feature extraction, and algorithms with two or more stages (Jarret *et al.*, 2009). The possibility of training neural networks with multiple intermediate layers fosters the emergence of several algorithms grouped in an area of study known as deep learning. The main objective of the algorithms that use two or more stages is to learn not only to distinguish between classes according to artificial descriptors, but also to be able to learn their own descriptors based on raw data. In the case of images, the learning is based on the pixel values themselves (Juraszek *et al.*, 2014).

Convolutional neural networks have been used for many years in image recognition, and they have obtained great success in recognizing characters in the research made by LeCun *et al.* (1990). Recent studies using convolutional neural networks known as Deep Convolutional Neural Networks (CNN) have reached the new state of the art in object recognition based in CIFAR-10 and NORB (Ciregan *et al.*, 2012). Generally, CNNs are trained under supervision, but research revealed that pre-training CNNs with filters obtained in a non-supervised way yields better results (Kavukcuoglu *et al.*, 2010).

In the research conducted by Sermanet *et al.* (Sermanet *et al.*, 2013), they present a framework using CNNs to perform recognition, localization, and detection of images. The contest winner was ImageNet Large Scale Visual Recognition Challenge 2013 (ILSVRC2013). The ILSVRC2013 database has 1.2 million images divided in 1,000 categories.

An important feature of CNNs is the possibility of reuse and fine tuning for different image bases. In the research conducted by Razavian *et al.* (2014), a pre-trained CNN named Overfeat was used to perform the extraction of a descriptor from different image bases whose CNN was not previously trained. In this case, the descriptors are then sorted by an SVM linear classifier. These results indicate a performance that is compatible with the state of the art, even when compared with algorithms that use manually-segmented images, a procedure that is unnecessary when CNNs are used and trained specifically in the analyzed basis.

Usage of deep learning is also described in the research made by Constante *et al.* (2016), who used a three-layer neural network with input through backpropagation. In this paper, this method was used to sort strawberries and obtained recognition results of 92.5% in the “Extra” category; 90% in the “Consumption” category; 90% in the “Raw material” category; and 100% in the “Alien objects” category.

3. Materials and Methods

The images of cattle branding presented in this research were provided by the City Hall, in Brazil. We used 39 cattle branding images. In the first experiment, we used 50 sample images per branding, totaling 1,950 images. For the second experiment, we used 70 sample images per branding, totaling 2,730 images. To all the sample images used in the experiments, we applied geometric transformations, such as size, scale and orientation variations. We intended to identify patterns as independently as possible from these variable factors. The images were provided in high resolution in the Portable Network Graphics format at a size of 253 x 253 pixels.

For the implementation of the proposed tool, for image database storage, algorithm processing, and viewing of the results, we used a personal computer with a video card that supports the CUDA parallel computing platform with a 5.0 compute capability version. Furthermore, we used the MATLAB software with the Neural Network, Parallel Computing and Statistics and Machine Learning libraries and the pre-trained convolutional network model obtained from the open source library VLFeat.org (VLFeat., 2016).

The proposed method consists of six steps, which are: image database selection; selection of pre-trained CNN model; pre-image processing and application of CNN; extraction of features of the images; training and classification of images by Support Vector Machines; and, finally, evaluation of the classification results. Fig. 1 illustrates a summarized flowchart of the proposed method.

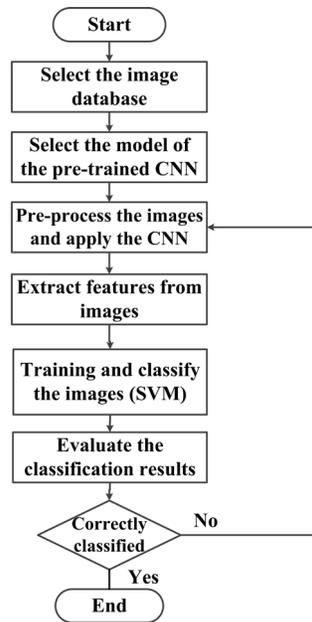


Figure 1: Summarized flowchart of the proposed method.

3.1. Image Database Selection

Fig. 2 illustrates some brands used in our proposed approach. The code of the brands obtained within the database provided by the City Hall in Brazil.

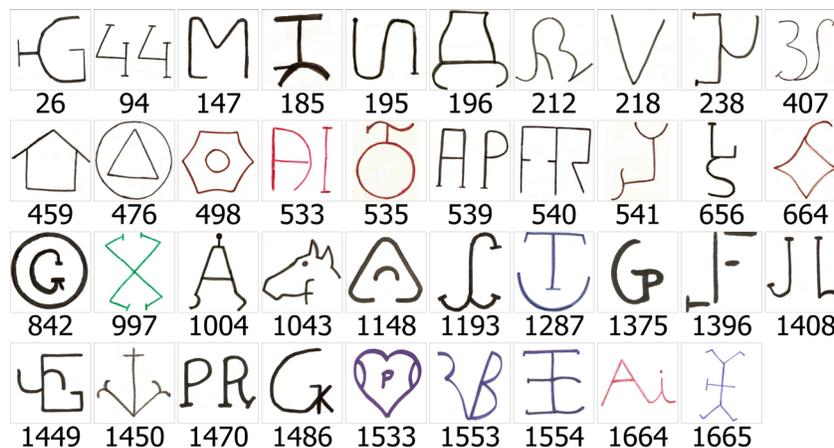


Figure 2: Cattle branding images used in our approach.

3.2. Model Selection of pre-trained CNN

Convolutional neural networks (CNN) are biologically-inspired architectures capable of being trained and of learning representations with no variations concerning scale, translation, rotation, and similar transformations (LeCun *et al.*, 2010). CNNs are one of the types of algorithms in an area known as deep learning. They are used with data in two dimensions, which make CNNs a good option for troubleshooting image recognition processes (Arel *et al.*, 2010). By definition, a deep architecture is a multiple-stage hierarchic framework, where each stage is composed by a neuronal network with at least 3 layers, and is trained through backpropagation.

Convolutional neural networks exhibit a relatively unexploited property known as knowledge transfer. This property is related to the fact that a CNN can be trained in an image base A (i.e., its weights are set for the classification of base A). Learning weights (and filters in CNNs) are considered generic enough to be used in the training of a new base B. By applying this concept, the convolutional neural network was pre-training using the VLFeat open source library. The use of the above mentioned pre-trained CNN model did not directly influence the recognition rate of cattle branding images.

3.3. Image Pre-processing and CNN Application

The adopted method is a neural network with five convolutional layers. Pre-processing is performed for gray-shaded images, with a triple replication of the images in order to create a RGB image.

The 1st convolutional layer has the 3 color channels as its input (RGB). Each convolution applies the non-linear activation function ReLU and a reduction through Maxpooling. The last layers are composed of fully connected neurons. Fig. 3 shows the proposed convolutional neural network architecture.

3.4. Feature Extraction

The set of filters learned by the CNN during training is responsible for detecting the features in the new image at the moment of a query. In the first filter level, we can observe some lines and orientations used for this detection. Fig. 4 illustrates the learned filters in the first convolutional layer using the RGB color space. There are 96 individual sets that represent the 96 filters used in that layer. It is possible to observe how the areas that present horizontal, vertical and diagonal bulges are highlighted after the first convolution.

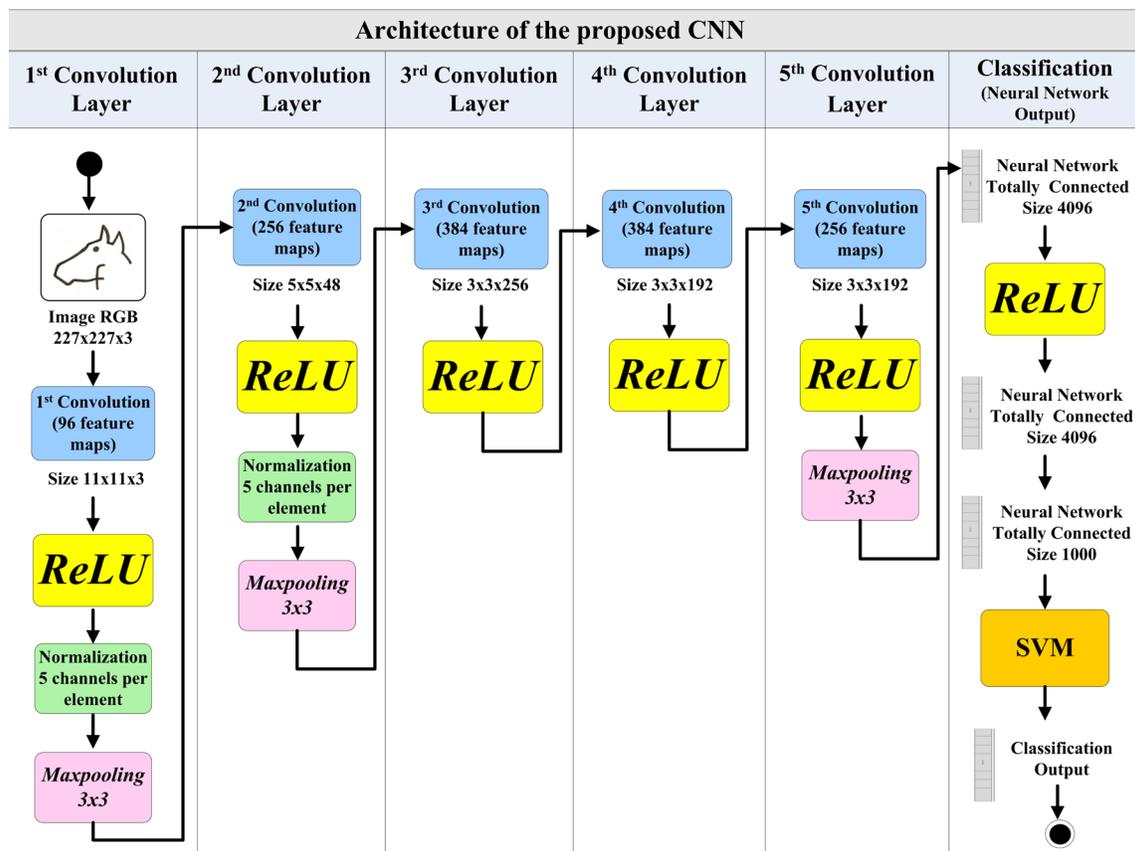


Figure 3: General model of a CNN architecture.

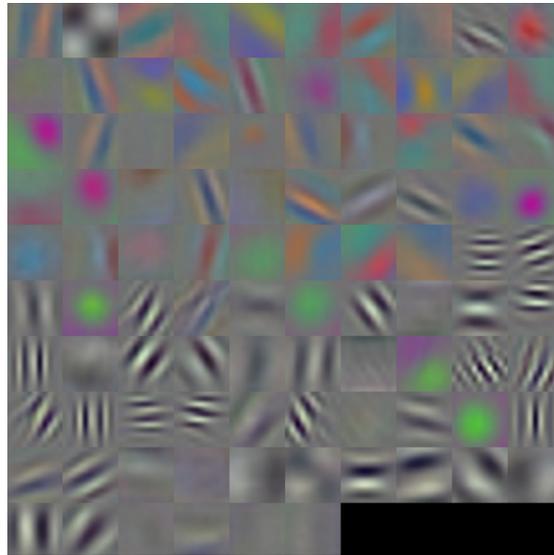


Figure 4: Convolutional 1st layer filters of the executed experiment.

After the first convolutions, it was possible to perform the extraction of image features for training the classifier. Fig. 5 presents an illustration of the algorithm for the extraction of features.

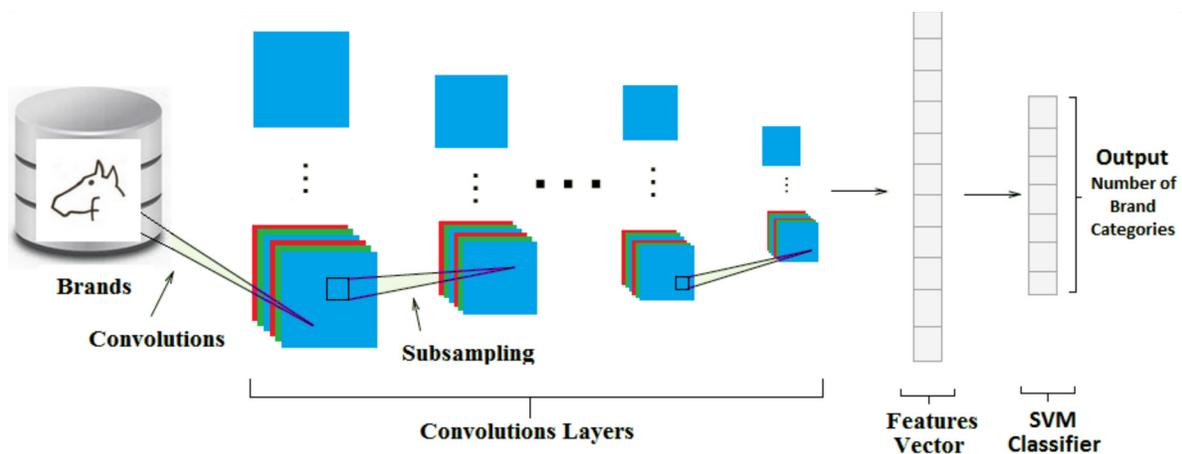


Figure 5: Illustration of the algorithm developed for the extraction of features.

3.5 Training and Classification of Images with Support Vector Machines

The model for automatic learning adopted in the presented work was the Support Vector Machine (SVM) supervised classifier. Support Vector Machine is a classification algorithm known for its success in a wide range of applications. SVMs are one of the most popular approaches for data modeling and classification. Its advantages include their outstanding capacity for generalization, which is highly related to the ability to correctly sort the samples that are not in the feature space used for training.

SVMs are used to classify and recognize patterns in several types of data; they are also employed in a wide range of applications, such as face recognition, clinical diagnosis, industrial process monitoring, and image processing and analysis (Tchangani *et al.*, 2005).

In regards to the tool proposed in this research, the classifier was used after the extraction of features from the brands belonging to distinct sample sets. During supervised learning, considering a series of examples (X_1, X_2), where X_1 represents an example and X_2 its classification, a classifier capable of foreseeing the class to which the new data belong, then performing the training process, should be produced. In the proposed research, we randomly separated the set of images in two parts, where one of these parts was used for the training stage and the other for the validation stage, thus eliminating polarization in the results. The final result is an average of the result obtained in the validation stage. The percentage division used here was 30% for training and 70% for validation.

3.6 Assessment of Classification Results

The confusion matrix contains information related to classifications performed through the application of a classifier. The performance of classifiers is often evaluated through data that are collected from this matrix. The confusion matrix of a classifier indicates the number of correct classifications versus the predictions made in each case over a group of examples. In this matrix, the lines depict the actual cases and the columns depict the predictions made by the model.

Through the confusion matrix, it is possible to find information related to the number of correctly and incorrectly classified images for each group of samples. This is an $A \times A$ matrix, where A is the amount of categories to which we apply the classifier. In our case, the conducted experiment included 39 brands. Thus, ours is a 39×39 confusion matrix in this situation. Based on the results obtained in confusion matrices generated by the conducted experiments, it was possible to apply metrics to the assessment of the proposed method in cattle branding recognition. The metrics used in this research are shown in Table 1.

Table 1: Metrics used in the evaluation of the proposed method

<i>Measure</i>	<i>Expression</i>
TPR	$TP / (TP + FN)$
PPV	$TP / (TP + FP)$
Acc	$(TP + TN) / (TP + FN + TN + FP)$
Err	$1 - \text{Acc}$
Kappa	$(P_o - P_e) / (1 - P_e)$

3. Results and Discussions

The proposed method made the assesment of the proposed method possible. The assessment of the results of the experiment was performed based on the recognition rate obtained in the confusion matrices generated from the classification attained in the validation stage of both experiments. Furthermore, the total processing time of the proposed method was also checked.

In the first experiment, we used 39 cattle brandings and a set containing 50 sub-images per branding; among those images, 30% were used for training. Fig. 6 presents the confusion matrix for the best result obtained in the first experiment, with a Overall Accuracy rate of up to 93%.

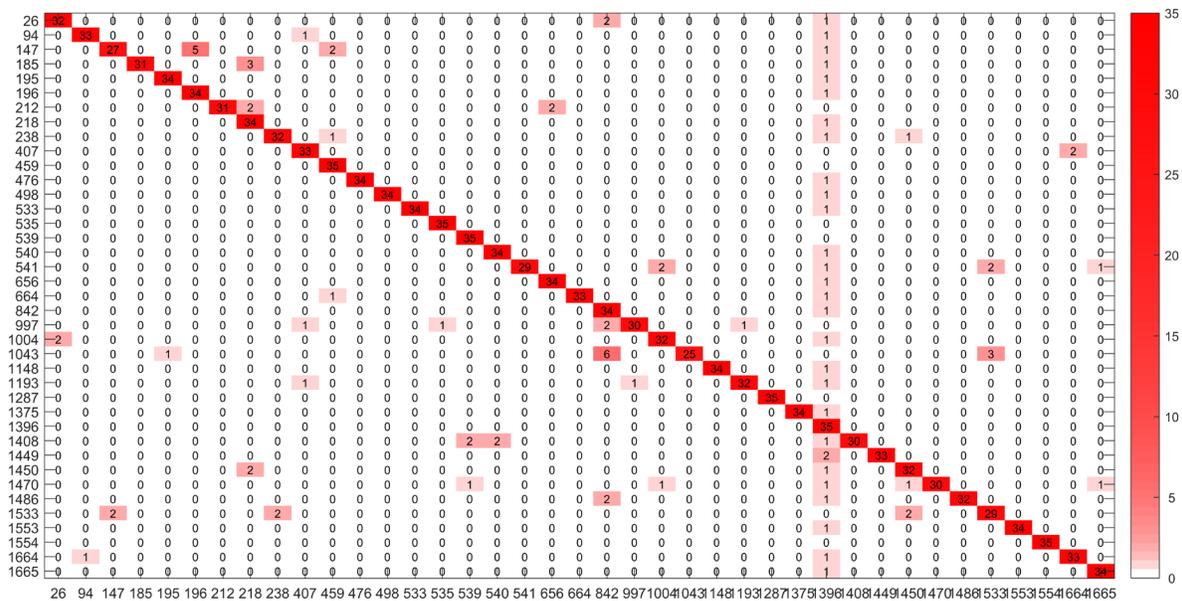


Figure 6: Confusion matrix of the First experiment – 50 sample images by brand.

Through the analysis of the main diagonal, we can observe that the correct classification rate is emphasized in five brandings: “459”, “535”, “539”, “1287”, “1396”, and “1554”, in which the percentage of correctness reached 100%. We can also observe that the branding with the lowest correct classification rates was “1043”, with a percentage of 71%.

The hypothesis of wrong classifications of cattle branding as shown in the confusion matrix is associated to the complexity of the samples, since some of the images present similar features among themselves. In general, the image samples of brandings with a better descriptive power and better quality correctly classified more images, since they include more characteristics when compared to brandings with worse sample quality, and, consequently, less features extracted. The capacity of recognizing patterns in an image based on a set of images depends on the amount of *a priori* information available about the object in question.

Another important fact to be presented, resulting from the analysis of Fig. 6, is the low number of false positives and false negatives, which are clearly located out of the main diagonal of the confusion matrix. This is due mainly to the capacity of the proposed method of extracting features from images, even in adverse scenarios, with images presenting different sizes, shapes, scales, orientations, noises, colors, and background contexts. In this kind of experiment, the noise in the images may impair the accuracy of the classification.

In the second experiment, we used 39 cattle brandings, but also using a set containing 70 sub-images per branding; among those images, 30% were used for training, differently from the 50 sub-images used in the previous experiment.

The goal of this experiment was to demonstrate that the increase in the number of sub-images used to extract features from a branding in the training phase contributes significantly to the accuracy performance of the Convolutional Neural Networks method. Fig. 7 presents the confusion matrix for the best result obtained in the experiment, with an Accuracy rate of up to 95%.

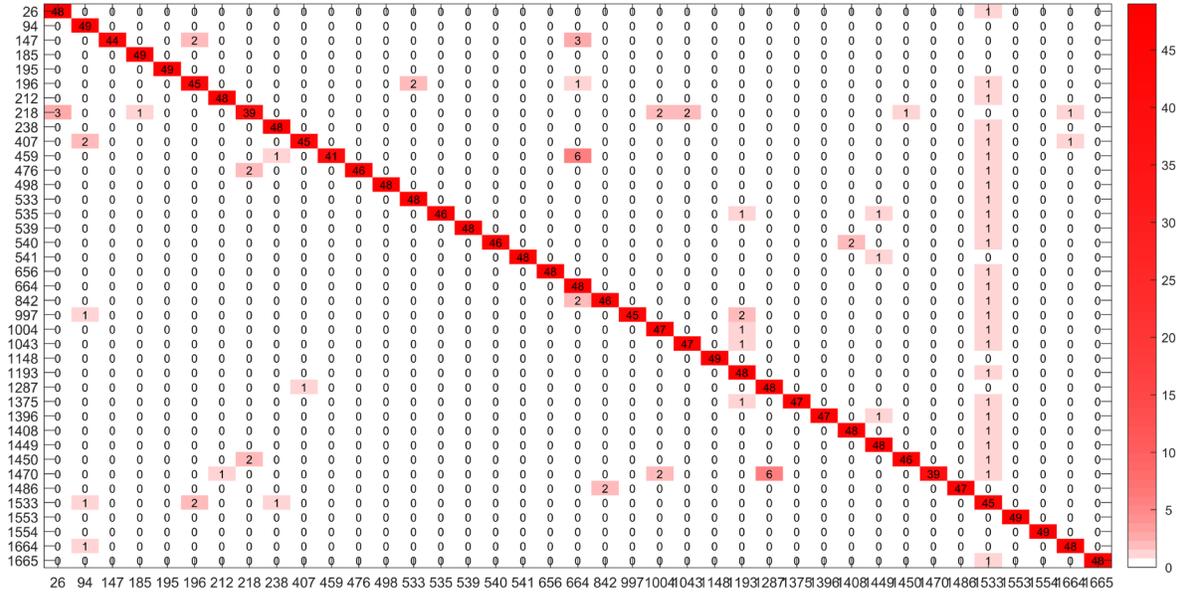


Figure 7: Second Experiment Confusion Matrix – 70 sample images per branding.

As for the second experiment, there was an improvement in the accuracy rate of 25 brandings when compared to the first experiment, or 64.10% from the total of brandings. This increase in the accuracy rates is explained by the increase in the number of sub-images for the representation of the brandings, which contributed significantly to the number of features extracted from the images that represent the brandings, increasing the correctness of the results and the percentage of overall accuracy.

We also assessed metrics related to image processing, such as Sensitivity (Recall) and Precision in order to define the reliability of the conducted experiments. Fig. 8 shows the Recall comparative chart (TPR) related to Experiments I and II, in which we can observe the best rates obtained in the second experiment.

Recall (TPR)

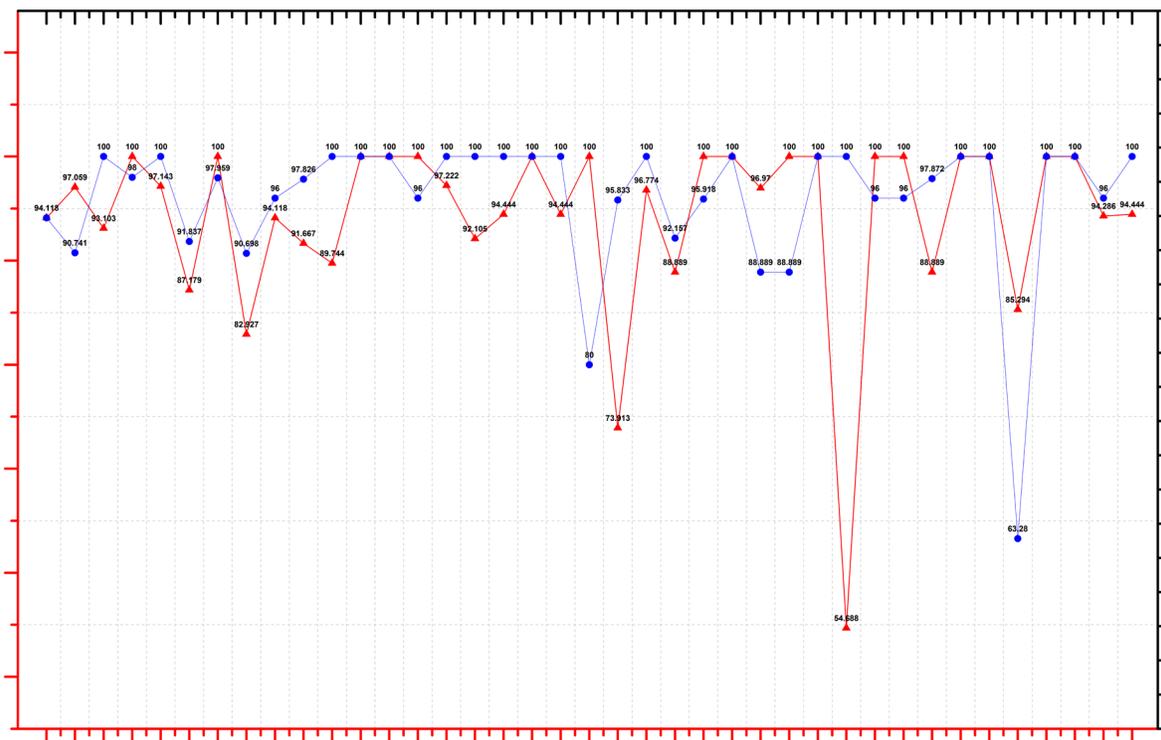


Figure 8: Compared recall – Experiments I and II.

Fig. 9 shows the Precision comparative chart (PPV) of Experiments I and II. By analyzing the chart, we can see that the Experiment II reached the best results, as well as in the Recall shown in Fig. 8.

4. Conclusion

In this research, we presented an automated method for the recognition of cattle branding. The project was developed and conducted in two institutions: Federal Institute of Farroupilha (IFFAR) and Federal Institute of Sul-Rio-Grandense (IFSUL). The experiments performed in this project used a Deep Learning technique of Convolutional Neural Network (CNN) to segmentation, as well as an SVM supervised classifier. In the CNN, we created a complete convolutional network, using images that were converted into the RGB color format as input. The experiments were all conducted using the cattle branding base provided by the City Hall in Brazil.

The proposed method achieved, for Experiment I, an accuracy of 93%, and an algorithm processing time of 32 seconds for the 39 assessed brands, in a total of 1,950 samples used for training and validation. In Experiment II, the method reached an accuracy of 95% and an algorithm processing time of 42 seconds for the same brands, in a total of 2,730 samples used for training and validation. The method proposed in this research showed better results regarding accuracy and processing time for the task of recognizing cattle brandings when compared the literature. Through the experiments, we found an increase in the percentage of correctness, as more sample images were added to the training and feature extraction phases.

For future researches, we intend to conduct experiments using CNN to classify cattle branding, in order to establish a comparison between the results attained through the application of the proposed hybrid method in this research, in which an SVM was used in the classification phase, and a neural network was used in the classification task.

5. References

- Atlas Socioeconômico, Brasil, <https://atlassocioeconomico.rs.gov.br/bovinos>, accessed 8 march 2019.
- R. Arnoni. Os Registros e Catálogos de Marcas de Gado da Região Platina. Pelotas: Revista Memória em Rede da UFPEL, 2013.
- G. Sanchez, M. Rodriguez. “Cattle Marks Recognition by Hu and Legendre Invariant Moments”. *ARPN Journal of Engineering and Applied Sciences*, vol. 11, N° 1, 2016.
- C. Silva, D. Welfer, F.P. Gioda, C. Dornelles. “Cattle Brand Recognition using Convolutional Neural Network and Support Vector Machines”. *IEEE Latin America Transactions*, vol. 15, N° 2, 2017. DOI: 10.1109/TLA.2017.7854627.
- X. X. Niu, C. Y. Suen. “A Novel Hybrid CNN-SVM Classifier for Recognizing Handwritten Digits”. *Pattern Recognition*, n. 45, p. 1318-1325, 2011. DOI: 10.1016/j.patcog.2011.09.021.
- K. Jarret, K. Kavukcuoglu, Y. LeCun. “What Is The Best Multi-Stage Architecture for Object Recognition?”. *IEEE 12th International Conference on Computer Vision*, p. 2146-2153, 2009. DOI: 10.1109/ICCV.2009.5459469.
- G. Juraszek. Reconhecimento de Produtos por Imagem Utilizando Palavras Visuais e Redes Neurais Convencionais, Joinville: UDESC, 2014.
- Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D Jackel. “Handwritten Digit Recognition with a Back-Propagation Network”. In: *Advances in Neural Information Processing Systems*. [S.l.]: Morgan Kaufmann, p. 396-404, 1990.
- D. Ciregan, U. Meier, J. Schmidhuber. “Multi-Column Deep Neural Networks for Image Classification”. In: *Proceedings of the 25th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012)*, [S.l.: s.n.]. p. 3642-3649, 2012. DOI: 10.1109/CVPR.2012.6248110.
- K. Kavukcuoglu, P. Sermanet, Y. Boreau, K. Gregor, M. Mathieu, Y. LeCun. “Learning Convolutional Feature Hierarchies for Visual Recognition”. In: *Advances in Neural Information Processing Systems*, ed. by J.D Lafferty and C.K.I. Williams and J. Shawe-Taylor and R.S. Zemel and A. Culotta, vol. 23, p. 1090-1098, 2010.
- P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun. “Overfeat: Integrated Recognition, Localization and Detection Using Convolutional Networks”, *CoRR*, abs/1312.6229, 2013.

- A. S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson. "CNN Features Off-the-Shelf: An Astounding Baseline for Recognition", In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, p. 806-813, 2014. DOI: 10.1109/CVPRW.2014.131.
- P. Constante, A. Gordón, O. Chang, E. Pruna, I. Escobar, F. Acuña. "Artificial Vision Techniques for Strawberry's Industrial Classification". IEEE Latin America Transactions, vol. 14, N° 6, 2016. DOI: 10.1109/TLA.2016.7555221.
- Vlfeat. Biblioteca Open Source VLFeat, <http://www.vlfeat.org/matconvnet/models/beta16/imagenet-caffe-alex.mat>, accessed 3 june 2016.
- Y. LeCun, K. Kavukcuoglu, C. Farabet. "Convolutional Networks and Applications in Vision". In: Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on. IEEE, p. 253-256, 2010. DOI: 10.1109/ISCAS.2010.5537907.
- I. Arel, D. Rose, T. Karnowski. "Deep Machine Learning - A New Frontier in Artificial Intellingence Research [research frontier]". Computational Intelligence Magazine, IEEE, v. 5, n. 4, p. 13-18. ISSN 1556-603X, 2010. DOI: 10.1109/MCI.2010.938364.
- A. Tchangani. "Support Vector Machines: A Tool for Pattern Recognition and Classification". Studies in Informatics & Control Journal, 14: 2. 99-109, 2005.
- J. Landis, G. Koch. "The Measurement of Observer Agreement for Categorical Data". International Biometric Society, v.33 n.1, p. 159, 1977. DOI: 10.2307/2529310.
- Y. Bengio, A. Courville, V. Vincent. "Representation learning: A review and new perspectives". IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 35, n. 8, p. 1798-1828, 2013. DOI: 10.1109/TPAMI.2013.50.
- G. Hinton. "To Recognize Shapes First Learn to Generate Images". Progress in Brain Research. Elsevier, v. 165, p. 535-547, 2007. DOI: 10.1016/S0079-6123(06)65034-6.
- M. Zeiler, R. Fergus. "Visualizing and Understanding Convolutional Networks". European Conference on Computer Vision. Springer, p. 818 - 833, 2014. DOI: 10.1007/978-3-319-10590-1_53.

