



Article

A New Way to Identify Mastitis in Cows Using Artificial Intelligence

Rodes Angelo Batista da Silva ^{1,*}, Héli-ton Pandorfi ^{1,*}, Filipe Rolim Cordeiro ², Rodrigo Gabriel Ferreira Soares ³, Victor Wanderley Costa de Medeiros ³, Gledson Luiz Pontes de Almeida ¹, José Antonio Delfino Barbosa Filho ⁴, Gabriel Thales Barboza Marinho ¹ and Marcos Vinícius da Silva ¹

¹ Department of Agricultural Engineering, Rural Federal University of Pernambuco, Recife 52171-900, PE, Brazil; rodesangel@gmail.com (R.A.B.d.S.); gledson.almeida@ufrpe.br (G.L.P.d.A.); gabriel.bmarinho@ufrpe.br (G.T.B.M.); marcos.viniciussilva@ufrpe.br (M.V.d.S.)

² Department of Computing, Rural Federal University of Pernambuco, Recife 52171-900, PE, Brazil; filipe.rolim@ufrpe.br

³ Department of Statistics and Informatic, Rural Federal University of Pernambuco, Recife 52171-900, PE, Brazil; rodrigo.gfsoares@ufrpe.br (R.G.F.S.); victor.wanderley@ufrpe.br (V.W.C.d.M.)

⁴ Department of Agricultural Engineering, Federal University of Ceará, Fortaleza 60020-181, CE, Brazil; zkdelino@ufc.br

* Correspondence: heliton.pandorfi@ufrpe.br

Abstract: Mastitis is a disease that is considered an obstacle in dairy farming. Some methods of diagnosing mastitis have been used effectively over the years, but with an associated relative cost that reduces the producer's profit. In this context, this sector needs tools that offer an early, safe, and non-invasive diagnosis and that direct the producer to apply resources to confirm the clinical picture, minimizing the cost of monitoring the herd. The objective of this study was to develop a predictive methodology based on sequential knowledge transfer for the automatic detection of bovine subclinical mastitis using computer vision. The image bank used in this research consisted of 165 images, each with a resolution of 360 × 360 pixels, sourced from a database of 55 animals diagnosed with subclinical mastitis, all of which were not exhibiting clinical symptoms at the time of imaging. The images utilized in the sequential learning transfer were those of MammoTherm, which is used for the detection of breast cancer in women. The optimized model demonstrated the most optimal network performance, achieving 92.1% accuracy, in comparison to the model with manual search (86.1%). The proposed predictive methodologies, based on knowledge transfer, were effective in accurately classifying the images. This significantly enhanced the automatic detection of both healthy animals and those diagnosed with subclinical mastitis using thermal images of the udders of dairy cows.

Keywords: image analysis; dairy cattle; convolutional neural network; infrared thermography



Citation: Silva, R.A.B.d.; Pandorfi, H.; Cordeiro, F.R.; Soares, R.G.F.; Medeiros, V.W.C.d.; Almeida, G.L.P.d.; Barbosa Filho, J.A.D.; Marinho, G.T.B.; Silva, M.V.d. A New Way to Identify Mastitis in Cows Using Artificial Intelligence. *AgriEngineering* **2024**, *6*, 4220–4232. <https://doi.org/10.3390/agriengineering6040237>

Academic Editor: In-Bok Lee

Received: 29 July 2024

Revised: 22 October 2024

Accepted: 30 October 2024

Published: 8 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bovine mastitis is the most devastating disease in the dairy industry worldwide. Detection and early diagnosis of bovine mastitis can play an important role in reducing economic losses resulting from delayed treatment of the disease. For this reason, the development of automated mastitis detection methods is increasingly encouraged [1–3].

Subclinical mastitis is an early stage of the disease with no visible signs [4]. It is a silent infection that manufacturers cannot detect with the naked eye, but which can be detected through on-site diagnostic methods such as somatic cell count (SCC) and microbial cultures [5].

The most common method is the California Mastitis Test (CMT), which is performed in the milking room as it is a cheap, quick, and easy procedure. However, it depends on human interpretation and can be less accurate and reliable [6]. In this context, developing

diagnostic methods that integrate automation is crucial to minimize costs and losses, ensure accuracy and speed in diagnosis, and promote non-invasive techniques like thermography to enhance animal welfare.

Measuring the surface temperature of the breast skin by infrared thermography is a noninvasive method capable of detecting temperature changes due to inflammation, used in the diagnosis of subclinical mastitis in dairy cows [1,4,7,8]. In these studies, the temperature intervals varied considerably, reaching a 5.3 °C difference between a healthy animal and those with subclinical mastitis.

When subclinical mastitis occurs, it has been shown that the local inflammatory response causes fever in the animal and changes in tissue blood flow [9,10], which can be the origin of an increase in the surface temperature of the udder skin [11]. The udder with mastitis has an elevated temperature even before clinical symptoms appear. In addition, Ref. [12] reported an increase of 2 to 3 °C in the udder surface temperature of lactating cows after inoculation with *Escherichia coli* in different parts of the udder. In a study carried out by the National Institute of Animal Health (NIAH), Ref. [13] observed that udder temperature measured using thermography can be a useful diagnostic tool for detecting mastitis in dairy cattle. Similarly, the authors [14] used thermographic images as a tool to detect mastitis in sheep, especially subclinical mastitis.

Screening for subclinical mastitis by measuring udder surface temperature has a high predictive diagnostic capacity, similar to the California Mastitis Test (CMT). However, analysis of the reliability of surface temperature by thermography among cows with different body and physiological characteristics living in different environmental conditions must be determined in each case [7].

In a study on dairy cows with mastitis [1], a software was developed that automatically measures udder temperature. The research demonstrated a strong correlation between these temperatures and the somatic cell count, which significantly facilitates the diagnosis of clinical mastitis. The findings indicate that elevated udder temperature is a reliable early indicator of mastitis, allowing for quicker intervention and better management of udder health. In another study, Ref. [15] compared automatic image recognition software with manual methods for detecting *E. coli* infections in cows. The results showed that the automated system provided comparable accuracy to manual methods, with the added benefits of increased speed and objectivity in identifying infections. This suggests that the implementation of image recognition technology could streamline the process of infection detection, improving both efficiency and reliability in herd management.

In a study based on convolutional networks [2], an object detector was developed using YOLOv3 as a proposal to perform detection and provide answers about the animal's clinical condition, achieving an accuracy of 83.33%. The transfer of convolutional network learning techniques was reported in a study [15], which was developed with digital images of nipples for the classification and identification of bovine mastitis, achieving an AUC of 0.920 or more. However, few studies have applied this technique to thermal images of animal udders with the main objective of obtaining the best model for their classification, since the use of optimized networks shows improvements in classification over the manual approach.

Among many deep learning models, convolutional neural network (CNN) is the most popular architecture. Due to its relatively simple and advanced features [16], it has been widely applied in the development of medical image analysis systems [17]. However, one of the main disadvantages of a convolutional neural network is that it requires a large amount of data to be trained, which is sometimes difficult to obtain in the field.

Therefore, transfer learning has recently become a popular strategy among researchers to solve this problem [17–21]. The current standard for transfer learning is to select and use existing models that were originally trained using large datasets of natural images, such as ImageNet [22], and then use limited images to model and adapt them.

Despite the advantages of CNNs, one of the main difficulties or challenges in using CNNs is training or fitting a model optimally. This is due to the fact that this can lead to the

problem of vanishing gradients, making it difficult to obtain the “best” or optimal values in the training process [23].

Therefore, many deep learning architectures have been developed and tested recently. A previous study compared several deep learning models and found that the ResNet50 model is the best architecture for image classification tasks with higher accuracy and training efficiency [24].

In this sense, machine learning algorithms have specific parameters known as hyperparameters. To adjust them, we need to choose the best model that fits the existing data. There are decision-theoretic methods for this, based on the concept of defining a search space of hyperparameters and determining the best combination of them [23].

Bayesian optimization is a hyperparameter search technique that surpasses both grid search and random search. Unlike these methods, Bayesian optimization leverages knowledge from previous iterations to inform the search process. This approach enhances decision-making by more effectively identifying the optimal hyperparameter settings for evaluating a model [25,26].

The objective of this study was to develop a prediction method based on sequential knowledge transfer using ResNet50 with ImageNet weighting and Bayesian hyperparameter optimization for the automatic detection of subclinical mastitis in cows by computer vision.

2. Materials and Methods

The transfer learning approach is a technique used to solve a given task using knowledge gained from solving a related task (Figure 1).

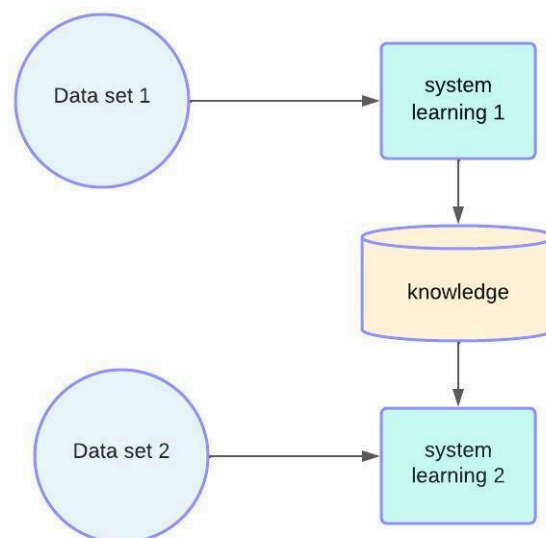


Figure 1. Transfer learning approach.

The deep network used in this approach utilized the ImageNet weights. This facilitates the learning of new features associated with new classes, allowing the network to start training optimally [27].

In order to utilize transfer learning, this study employed a sequential transfer model with a pre-trained network for the purpose of identifying thermal images of the udder of dairy cows. A pre-trained neural network fine-tuning technique was utilized for the creation of a learning transfer model, the principal function of which is the transfer of information between related domains.

The applications used in this study were characteristic extraction and transfer learning as classifiers. In this case, in a layered architecture, a deep learning model learns different features in multiple layers [28].

This layered architecture allows ResNet50 not to have a final layer (FC), using convolutional layers as fixed feature extractors and FC layers adapted to the classes used (healthy and sick animals). You can use pre-trained networks [28].

2.1. Sequential Transfer Approach

The deep network used in this approach utilizes the ImageNet weights. This facilitates the learning of new features associated with new classes and improves the generalization of the network [29].

For the mastitis classification problem, the network was pre-trained on the first domain of the ImageNet database (this dataset consists of 1000 object classes and contains 1,281,167 training images, 50,000 validation images, and 100,000 test images) and then adapted to human mammography images. Classification task (second domain). Finally, we adjusted the presented network and trained it to classify thermal images of bovine mastitis (the target domain) (Figure 2).

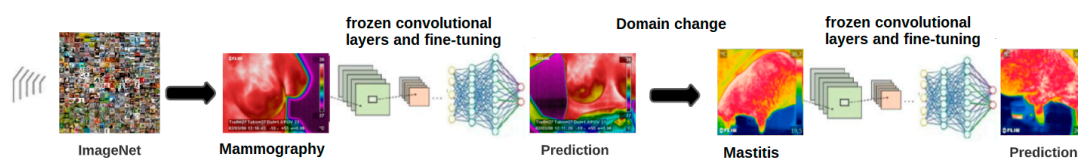


Figure 2. The general procedure of sequential transfer learning.

2.2. Optimization of Hyperparameters

Machine learning algorithms have certain parameters, also known as hyperparameters, so we need to choose the optimal setting. A decision-theoretic approach may be employed to address this issue. This entails defining a search space of hyperparameters and determining the optimal combination thereof [23].

The approaches used to adjust the hyperparameters in this study included manual search and Bayesian optimization. In manual search, hyperparameter tuning consisted of model selection based on previous experience. Then, a model was trained and evaluated using these parameters. This process was repeated for different sets of values until maximum accuracy was achieved or the model reached an optimal error.

Bayesian optimization is more accurate and faster than manual approaches. The best model parameters are selected automatically.

The problem with using a manual approach is that adjusting these parameters can be time-consuming and affect the final performance of the network. However, Bayesian techniques make it easier to approximate the perfect solution. Bayesian optimization enhances decision-making about which hyperparameter setting is best for evaluating a model [25,26].

2.3. Data Set

The ImageNet dataset is based on a dataset of over 15 million high-resolution tagged images from approximately 22,000 categories. The images were collected from the web and tagged by markers using Amazon's Mechanical Turk crowdsourcing tool [30].

To train the network and compare the model, we used a larger database with a similar problem (detection of breast cancer in women) using thermal images. These images are part of the MammoTherm database of the Federal University of Pernambuco. They have 640×480 pixels and are being used at the Hospital das Clínicas of the same institution. This image database is public and has been made available for research on breast cancer in women.

The training set, the test set, and the validation set consisted of 600 images from the database (MammoTherm), considering images from the groups "No lesion" and "Benign lesion" [31,32].

The results were subsequently transferred to a smaller database (images of healthy cows and cows with subclinical mastitis). The percentage of images used in the training,

testing, and validation phases was 70, 15, and 15%, respectively, for the mammography image database, and the same percentages for the bovine mastitis image database.

The image bank used in this study comes from a preliminary study approved by the Ethics Committee on the Use of Animals of the Federal Rural University of Pernambuco 138/2018, approved on 19 December 2018, and created from a mammary thermography record of animals with clinical diagnosis, subclinical mastitis, and free of mastitis (healthy). Thermal images of the udders of dairy cows from three locations (latitude: 8°36'33" S, 36°37'30" W, altitude 733° m) were obtained in different production units of the city of Capoeiras. Additional data were collected from Pesqueira (8°21'35" S, 36°41'42" W, 652 m) in the State of Pernambuco, Brazil, provided by NEAMBE, Center for Studies of Agricultural Atmosphere and Animal Welfare at the Federal University of Ceará.

The number of samples was determined based on the selection criteria of animals (Giroland cows) of the same calving order, lactation stage, body score, milk production, and quality. The clinical status of the animals was assessed using the California Mastitis Test (CMT) before milking and after discarding the first drop of milk. The test was carried out for each mammary quarter, with scores ranging from 0 to +++, whereby at score zero there was no precipitate formation (healthy), at score traits there was mild precipitation (trace infection), in score + there was moderate precipitation (subclinical mastitis), in score ++ there was clear gel formation (subclinical mastitis), and in score +++ there was marked gel formation (subclinical mastitis). To limit subjectivity in interpreting the results, only those with scores between + and +++ were considered for selecting animals with subclinical mastitis, and the sensitivity range for detecting sick animals was 93% for SCC > 500,000 cells/mL [33] (Figure 3).

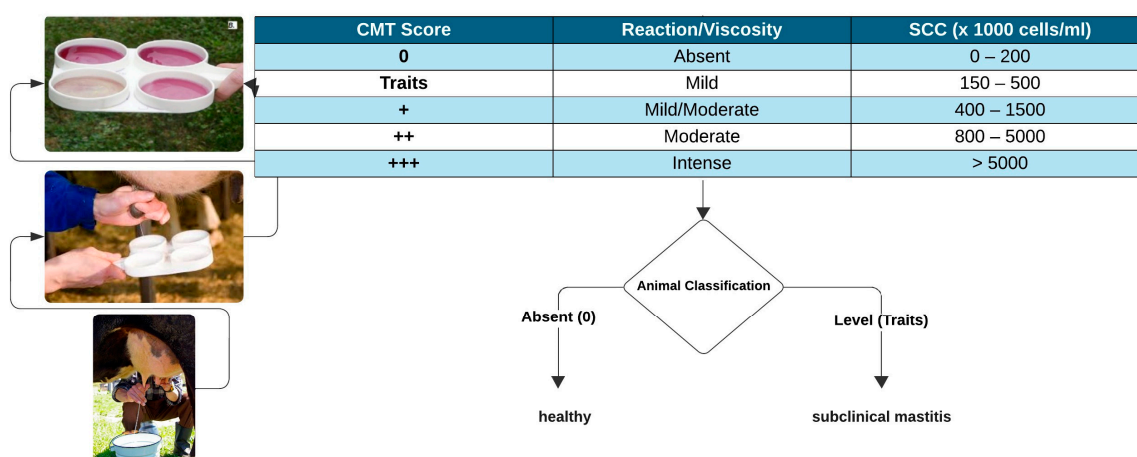


Figure 3. Identification of subclinical mastitis by CMT and classification of healthy animals and those with subclinical mastitis.

Images were acquired using a FLIR i60 thermal imaging camera with a focal length of 1 m and emissivity (ϵ) of 0.98 specified for biological tissue [8]. Images were obtained between 05:00 and 07:00 of the morning shift, before the first milking of the animal, including left anterolateral, right anterolateral, and posterior images, 3 images per animal (Figure 4).

A total of 165 labeled images, each with a resolution of 360×360 pixels, were extracted from a database comprising 55 cattle. These images were classified into two distinct groups: 'Healthy' and "Subclinical Mastitis", based on CMT, which were used as training, testing, and validation data sets.

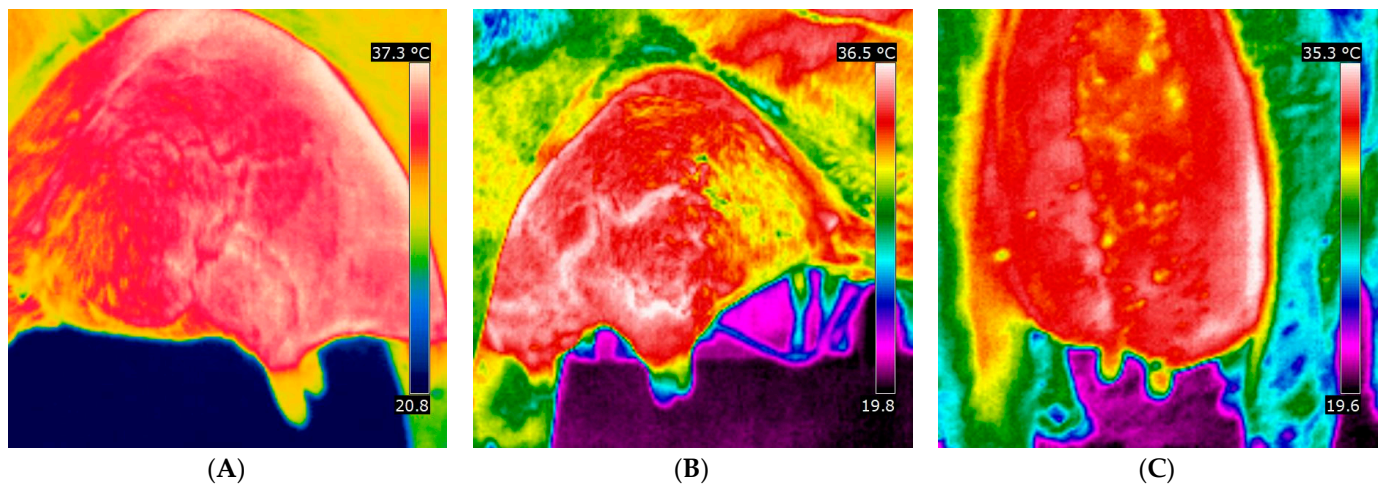


Figure 4. Thermal images of the right (A), left (B), and posterior (C) anterolateral frames.

2.4. Image Preprocessing

In the pre-processing, we resized all the udder images of dairy cattle to 224×224 pixels and used a data augmentation technique to enlarge the dataset. The same number of images was used for each class in all image banks. For this, some transformations were performed only once to generate surplus images for the network. This is commonly used to avoid overfitting and increase the strength of the system [15].

In machine learning, the performance of the algorithm can be improved depending on the amount of data available. This almost never happens. Data Augmentation (DA) uses computational techniques to augment a supervised training set to obtain an algorithm with higher predictive power and consequently better results [34].

In this step, a geometry method was used, which consisted of transformations that changed the shape of the image by mapping individual pixel values to new destinations. The basic shape of the layer represented in the image is maintained but modified to new positions and orientations. Rotation of 30° , magnification of 20° , cropping of 10° , and random horizontal and vertical displacement of 20° were used, and 1500 new images were created (Figure 5).

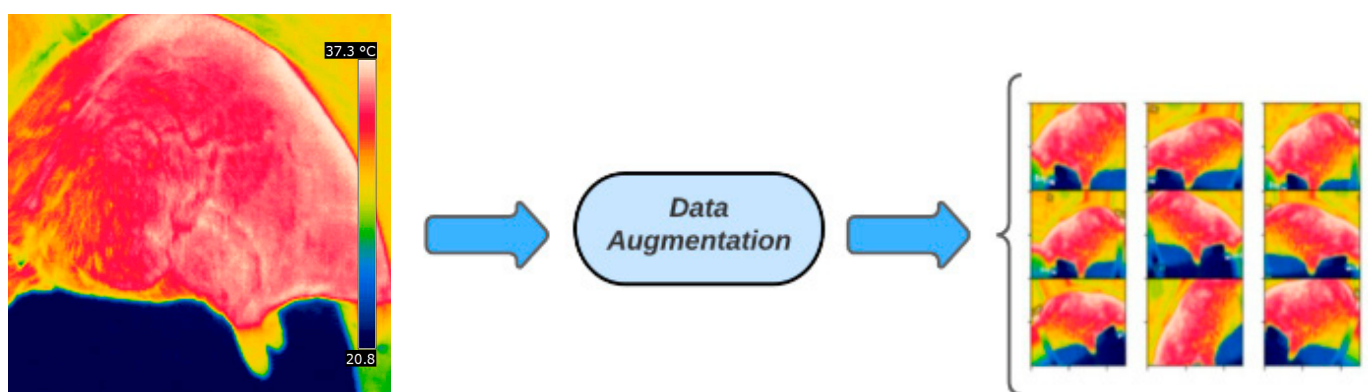


Figure 5. Image results after using data augmentation.

2.5. Hyperparameter Fitting Using Bayesian Optimization

An action that balances the choice between exploring the entire search space and exploring strong regions of the search space. Figure 6 shows the optimization process.

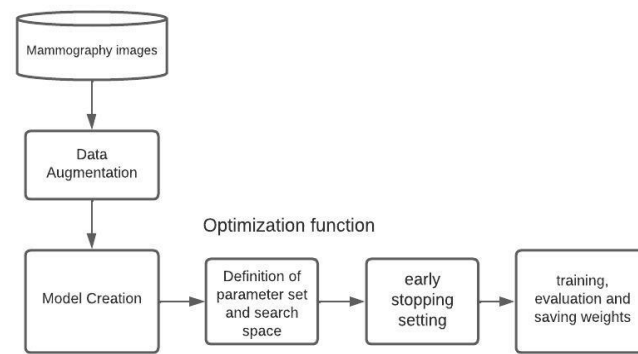


Figure 6. Bayesian optimization in the developed system.

Bayesian optimization treats parameters as random variables associated with probability distributions. In this case, an observation is the complete training of the CNN model with the hyperparameters chosen for that observation. For each iteration, a set of hyperparameters was selected and observations were made. The validation accuracy was used to evaluate the observations. The set of hyperparameters was chosen based on the detection bias. In this case, an observation is the complete training of the CNN model with the hyperparameters chosen for that observation. For each iteration, a set of hyperparameters was selected and observations were made. A Gaussian process with radial basis function cores was used, such that the model is optimized based on a single accuracy measure on the validation dataset.

Bayesian optimization implemented in the Python GPyOpt 1.2.6 library was used to develop the system. This allows for different parameters that affect the performance of the optimization. The default parameters provided by the library were used because of the interest in the possible performance improvement in combination with sequential transfer learning.

This approach was used in two phases, with two sets of hyperparameters, one for each training phase. Two separate Bayesian optimization runs were performed. First, the classification block was optimized for mammography images and then the best model was used in the second step.

The hyperparameter tuning procedures target some of the most important parameters such as the number of units in densely connected layers, the learning rate, and the activation function [28]. We added other important network regularization parameters such as the l2 parameter and the dropout rate for densely connected layers using the RMSprop optimizer (Table 1).

Table 1. Hyperparameters and search space.

Parameter	Search Space	Description
Number of neurons in the custom dense layer	64, 128, 256, 512	To introduce nonlinearity in the output of the neurons
Learning Rate	0.01; 0.001; 0.0001	To determine the step size in each iteration while minimizing the loss function
l2	0.1; 0.01; 0.001	Regularization
Activation Function	Relu, Elu, and Tanh	To introduce nonlinearity in the output of neurons
Dropout rate	0.3; 0.5; 0.7	To ensure that the model is robust to the loss of any individual evidence

2.6. Evaluation Metrics

Following the same performance metrics in the references related by [35–37], the network performance was evaluated based on the following metrics: accuracy (Equation (1)), F1 score (Equation (2)), and visualization by the confusion matrix.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (2)$$

where: *TP*—True Positive; *TN*—True Negative; *FP*—False Positive; *FN*—False Negative.

3. Results

In this section, the performance of the network using Bayesian optimization on the task of mastitis image classification is presented. The accuracy and loss curves for the training and test classification results of the images obtained from the transfer learning networks are presented in Figure 7.

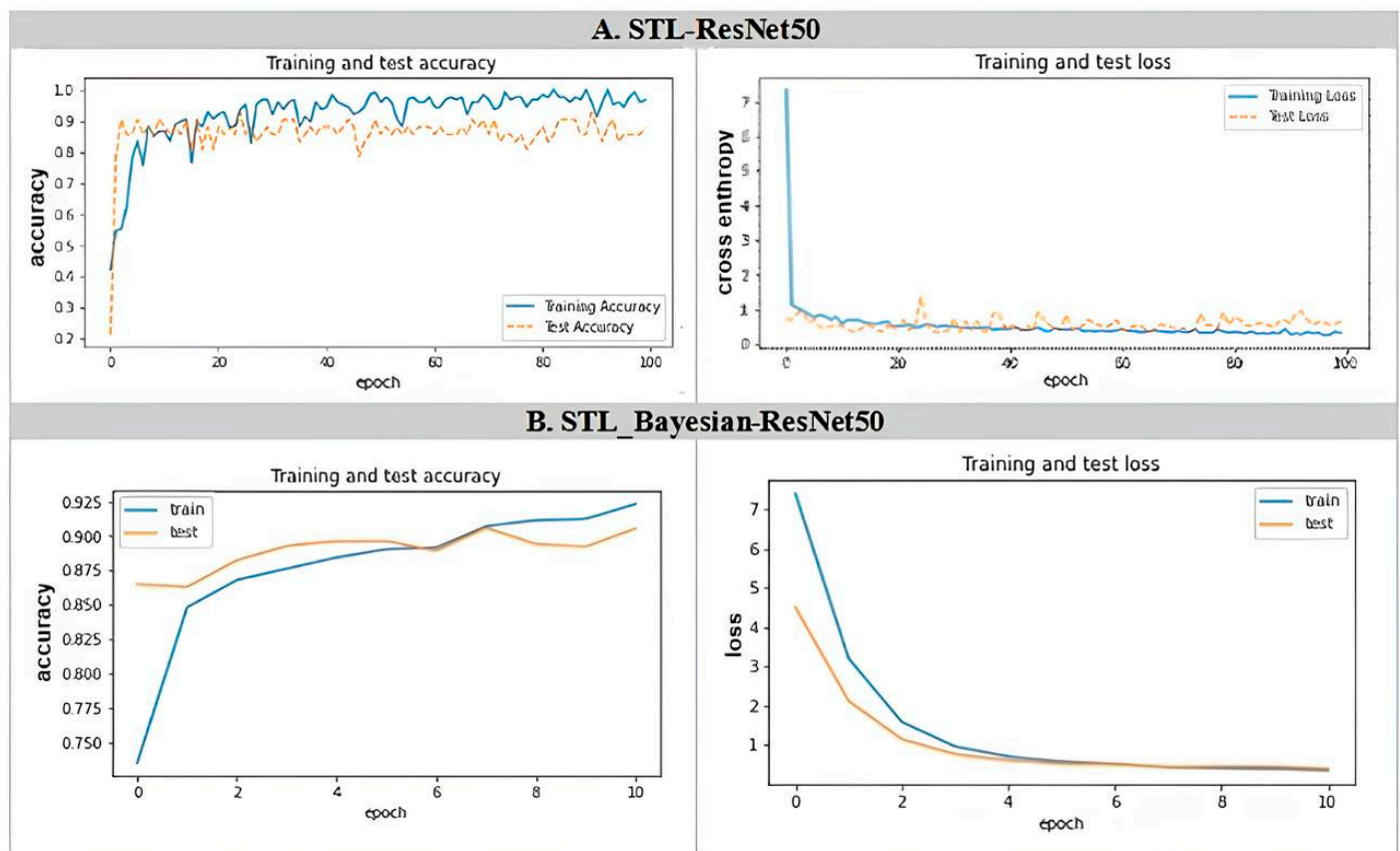


Figure 7. Accuracy and loss curves in training and testing for the fitted models.

The optimized model exhibited the most optimal network performance, achieving a 92.1% accuracy in comparison to the model that utilized a manual search (88.03%).

3.1. Classification Performance with Bayesian Optimization

By obtaining the optimization plots (Figure 8), it is possible to visualize the exploratory space of the Bayesian model, which starts by testing a wide range of distant hyperparameters in search of the best performance, with the most satisfactory accuracy range.

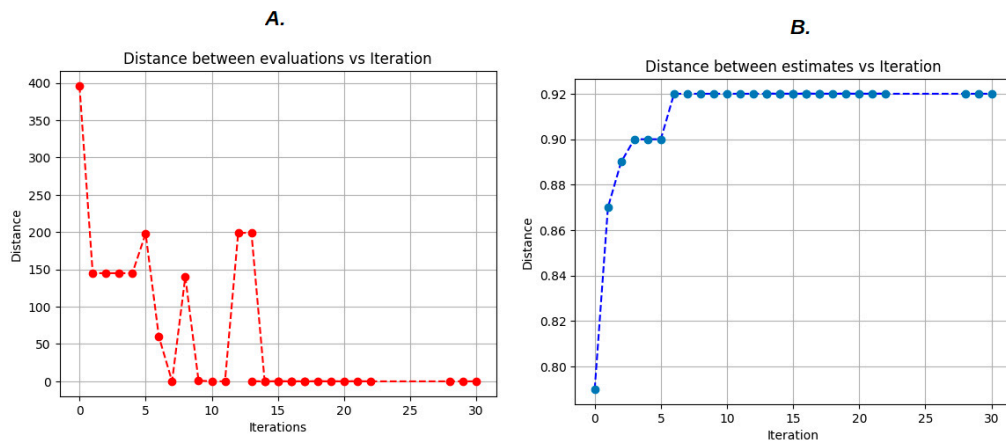


Figure 8. Bayesian optimization plots: (A) Distance between consecutive trials, (B) Precision vs. iteration.

Figure 8A shows how convergence to the minimum value of an objective function is performed with respect to the number of different attempts of hyperparameter combinations. The x-axis shows the number of calls to the objective function, and the y-axis shows the minimum value of the objective function after several calls. There was less distance in the evaluations from the thirteenth iteration, as the evaluations were very close, showing that the network found the optimal values of the hyperparameters, ending the training. The same did not occur between some iterations, such as the sixth and tenth, as they presented a greater distance between them (Figure 8A).

3.2. Model Performance in Classification and Set of Hyperparameters

The performance of the models in the task of classification of thermal images of the udder of dairy cows (healthy animals and animals with subclinical mastitis) was compared using the visualization of the confusion matrix (Figure 9).

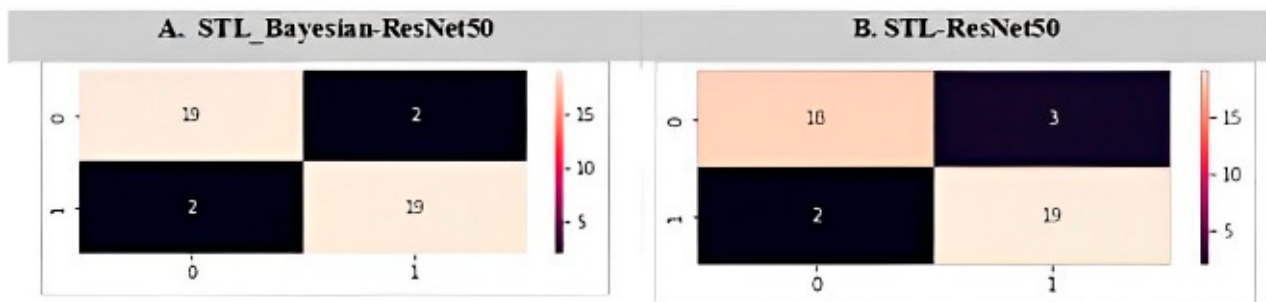


Figure 9. Confusion matrices of the four models employed in the classification of the images (0—healthy animal; 1—an animal with subclinical mastitis).

Based on the results of the confusion matrix showing the number of correctly classified images, the model had 19 correctly classified images and two incorrectly classified images for the two classes of healthy animals and animals with subclinical mastitis.

The best set of hyperparameters and best set of weights for training the network conferred from the Bayesian surrogate model are shown in Table 2.

The number of neurons in the customized dense layer was 64 and the learning rate for training on the image set was 0.0001. The rate used in l2 regularization was 0.001, with an activation function of elu, and a dropout rate of 0.5.

Table 2. The best set of hyperparameters inferred by the Bayesian surrogate model.

Parameter	Search Space
Number of neurons in the custom dense layer	64
Learning Rate	0.0001
l2	0.001
Activation Function	Elu
Dropout	0.5

4. Discussion

ImageNet weights were used in the system because they are more efficient in transmitting relatively small datasets, speeding up the technique compared to training from scratch. The images used are different from those in the ImageNet domain but resemble natural images in terms of color, contrast, and lighting. These common image features were important for classification. This latter accuracy was obtained in 100 seasons, which did not occur with the optimized model, with the best performance obtained in 11 seasons.

The F1 scores showed values of 90 and 85.5%. This comprehensively considers accuracy and recovery, with the results being able to better reflect the superior performance of the model along with accuracy. This is because reducing false negatives helps with health management, increases confidence in diagnosis, reduces costs, and ensures a higher value for the milk paid to producers. In general, increased accuracy allows producers to act more quickly to treat subclinical mastitis and prevent it from progressing to a more chronic stage, maintaining milk production, reducing waste, and improving the health and well-being of the herd.

As can be seen in Figure 8A, the distances represented in the graph explain which regions represent the most accurately punctuated search space. This means that these areas can provide better results and be closer to the ideal. In this case, the results displayed after the 10th attempt allowed the algorithm to reach the highest accuracy value.

The precision values of the 34 trials performed by the algorithm remained constant and were above 90% between the 2nd and 22nd iterations. After that, the algorithm reached the best value (92.1%) without changing anything in the next run (Figure 8B).

Therefore, the algorithm converged quickly to a set of hyperparameters and achieved high accuracy with only small incremental improvements. These results show the efficiency with which the Bayesian model can converge to identify the optimal set of hyperparameters for the main classifier based on the pre-trained CNN ResNet50.

Compared to the model without Bayesian optimization, it presented 18 correctly classified images for healthy animal classes, resulting in better performance, but the same percentage for animals with subclinical mastitis.

In Figure 9A, STL_Bayesian—ResNet50 has a higher identification rate (true positive rate and true negative rate). The identification rate reached 92% in healthy animals and animals with subclinical mastitis. The results of this study showed that the ability to detect subclinical mastitis was superior to the research carried out by [2], where the accuracy of the mastitis classification algorithm was 83.33%. This is mainly due to the use of fine-tuning as an optimization of hyperparameters in the training images. We first adapted the network to the new class used and then adapted the network to the mastitis image to achieve good performance. Similar results were achieved by the study carried out by [38] to detect subclinical mastitis using a deep learning model based on Convolutional Neural Networks (CNN) using 7615 udder thermograms from 40 Murrah buffaloes and had training accuracy and validation accuracy of 0.970 and 0.943, respectively. Thus, the improved deep learning CNN models efficiently predicted cases of subclinical mastitis.

A higher percentage of errors (false positives and false negatives) was observed in the STL—ResNet50 model (14% and 8%) (Figure 9B). This is mainly due to manual fitting, so the values of the selected hyperparameters affect the quality of models, as well as their ability to learn and generalize.

In this context, it is possible to infer that infrared thermography was a technique capable of mapping the udder surface temperature of dairy cattle, identifying changes resulting from the clinical picture of bovine mastitis. When combined with image analysis techniques, it enabled the accurate diagnosis of subclinical mastitis through deep machine learning, minimizing the cost of having to test the entire herd through CMT.

Deep learning and sequential knowledge transfer were used to improve the classification accuracy of thermal images associated with the clinical picture of dairy cattle, which contributed significantly to the automatic detection of healthy animals and animals with subclinical mastitis. This predictive methodology contributes significantly to reducing false positives by only referring animals with an alert indication for subclinical mastitis to the CMT test.

The use of Bayesian optimization, whose motivation is linked to the choice of the best parameters for a deep learning model, interfered with the performance of the models. Compared to the manual choice of parameters, it represents a time saver because the manual choice is laborious. The mechanisms of attention applied to computer vision tasks are inspired by human visual capacities, more precisely attention, which consists of focusing on a certain object in a scene. When applied to deep learning models, they improve their performance by highlighting discriminative characteristics in the image, helping with classification.

The results of this study indicate that image processing techniques applied to thermal images were able to extract characteristics that helped classify the images; however, more studies need to be carried out involving convolution networks and their techniques in order to obtain new results that help producers make decisions. In terms of detection time, the proposed predictive methodology is faster and more objective, directing animals to the CMT test. In relation to SCC, Electronic Somatic Cell Count, and other laboratory methods (such as microbial culture), the proposed methodology can also support the identification of animals that will have their milk samples tested.

5. Conclusions

The predictive model STL_Bayesian—ResNet50 correctly classified the images, obtaining an accuracy of 92.1%.

Compared to the STL model, ResNet50 achieved a 6% higher performance and converged in a significantly lower number of epochs (11 epochs).

The implementation of Bayesian optimization enhanced the functionality of the network, identifying the optimal set of parameters that significantly contributed to the automated detection of healthy animals and those exhibiting subclinical mastitis from thermal images of the udder of dairy cows.

Follow-Up Research Plan

The results of the present study indicate that the processing techniques of images applied to thermal images aiming at the early diagnosis of mastitis were able to extract characteristics that helped in the classification of images. However, more studies are needed involving convolutional networks and their techniques in order to obtain new results that assist the producer in decision-making.

Author Contributions: Conceptualization, H.P. and R.A.B.d.S.; methodology, F.R.C. and R.A.B.d.S.; software, R.A.B.d.S.; validation and formal analysis, R.A.B.d.S., R.G.F.S. and V.W.C.d.M.; research, H.P., G.L.P.d.A. and J.A.D.B.F.; resources, H.P. and G.L.P.d.A.; data curation, R.A.B.d.S.; writing—preparation of the original draft, R.A.B.d.S.; writing—proofreading and editing, R.A.B.d.S. and H.P., M.V.d.S. and G.T.B.M.; visualization, R.A.B.d.S., F.R.C., G.L.P.d.A. and J.A.D.B.F.; supervision, H.P.; project administration, H.P., G.L.P.d.A. and R.A.B.d.S.; obtaining funding, H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Federal University of Pernambuco—UFRPE and the Foundation for the Support of Science and Technology of the State of Pernambuco—FACEPE-APQ-0496-5.03/22.

Institutional Review Board Statement: The image bank used in this research comes from a preliminary study approved by the Ethics Committee on Animal Use of the Federal Rural University of Pernambuco, license No. 138/2018, approved on 19 December 2018, formed from the registration of thermal images of the udder of animals with a previous diagnosis of clinical, subclinical, and healthy mastitis.

Data Availability Statement: All the data that support the findings of this study are available from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zaninelli, M.; Redaelli, V.; Luzi, F.; Bronzo, V.; Mitchell, M.; Dell'Orto, V.; Bontempo, V.; Cattaneo, D.; Savoini, G. First Evaluation of Infrared Thermography as a Tool for the Monitoring of Udder Health Status in Farms of Dairy Cows. *Sensors* **2018**, *18*, 862. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Zhang, X.; Kang, X.; Feng, N.; Liu, G. Automatic recognition of dairy cow mastitis from thermal images by a deep learning detector. *Comput. Electron. Agric.* **2020**, *178*, 105754. [\[CrossRef\]](#)
3. da Silva, R.A.B.; do Monte Lima, J.P.S.; Pandorfi, H.; de Almeida, G.L.P. Thermal Image Thresholding for Automatic Detection of Bovine Mastitis. *Int. J. Comput. Appl.* **2021**, *183*, 29–33. [\[CrossRef\]](#)
4. Watz, S.; Petzl, W.; Zerbe, H.; Rieger, A.; Glas, A.; Schröter, W.; Landgraf, T.; Metzner, M. Technical note: Automatic evaluation of infrared thermal images by computerized active shape modeling of bovine udders challenged with *Escherichia coli*. *J. Dairy Sci.* **2019**, *102*, 4541–4545. [\[CrossRef\]](#)
5. Viguiet, C.; Arora, S.; Gilmartin, N.; Welbeck, K.; O'Kennedy, R. Mastitis detection: Current trends and future perspectives. *Trends Biotechnol.* **2009**, *27*, 486–493. [\[CrossRef\]](#)
6. Ruegg, P.L. A 100-Year Review: Mastitis detection, management, and prevention. *J. Dairy Sci.* **2017**, *100*, 10381–10397. [\[CrossRef\]](#)
7. Polat, B.; Colak, A.; Cengiz, M.; Yanmaz, L.E.; Oral, H.; Bastan, A.; Kaya, S.; Hayirli, A. Sensitivity and specificity of infrared thermography in detection of subclinical mastitis in dairy cows. *J. Dairy Sci.* **2010**, *93*, 3525–3532. [\[CrossRef\]](#)
8. Da Silva, R.A.B.; Pandorfi, H.; Almeida, G.L.P.; De Assunção Montenegro, A.A.; Da Silva, M.V. Spatial dependence of udder surface temperature variation in dairy cows with healthy status and mastitis. *Rev. Bras. De Saúde E Produção Anim.* **2019**, *20*, e01102019. [\[CrossRef\]](#)
9. McManus, R.; Boden, L.A.; Weir, W.; Viora, L.; Barker, R.; Kim, Y.; McBride, P.; Yang, S. Thermography for disease detection in livestock: A scoping review. *Front. Vet. Sci.* **2022**, *9*, 965622. [\[CrossRef\]](#)
10. Yang, C.; Li, G.; Zhang, X.; Gu, X. Udder skin surface temperature variation pre- and post- milking in dairy cows as determined by infrared thermography. *J. Dairy Res.* **2018**, *85*, 201–203. [\[CrossRef\]](#)
11. Bortolami, A.; Fiore, E.; Giancesella, M.; Corro, M.; Catania, S.; Morgante, M. Evaluation of the udder health status in subclinical mastitis affected dairy cows through bacteriological culture, somatic cell count and thermographic imaging. *Pol. J. Vet. Sci.* **2015**, *18*, 799–805. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Pezeshki, A.; Stordeur, P.; Wallemacq, H.; Schynts, F.; Stevens, M.; Boutet, P.; Peelman, L.J.; De Spiegeleer, B.; Duchateau, L.; Bureau, F.; et al. Variation of inflammatory dynamics and mediators in primiparous cows after intramammary challenge with *Escherichia coli*. *Vet. Res.* **2011**, *42*, 15. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Rodrigues, A.L.; de Santana, M.A.; Azevedo, W.W.; Bezerra, R.S.; Barbosa, V.A.F.; de Lima, R.C.F.; dos Santos, W.P. Identification of mammary lesions in thermographic images: Feature selection study using genetic algorithms and particle swarm optimization. *Res. Biomed. Eng.* **2019**, *35*, 213–222. [\[CrossRef\]](#)
14. Berry, R.J.; Kennedy, A.D.; Scott, S.L.; Kyle, B.L.; Schaefer, A.L. Daily variation in the udder surface temperature of dairy cows measured by infrared thermography: Potential for mastitis detection. *Can. J. Anim. Sci.* **2003**, *83*, 687–693. [\[CrossRef\]](#)
15. Wilson, A.C.; Roelofs, R.; Stern, M.; Srebro, N.; Recht, B. The marginal value of adaptive gradient methods in machine learning. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4148–4158.
16. Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **2018**, *172*, 1122–1131. [\[CrossRef\]](#)
17. Porter, I.R.; Wieland, M.; Basran, P.S. Feasibility of the use of deep learning classification of teat-end condition in Holstein cattle. *J. Dairy Sci.* **2021**, *104*, 4529–4536. [\[CrossRef\]](#)
18. Boumaraf, S.; Liu, X.; Zheng, Z.; Ma, X.; Ferkous, C. A new transfer learning-based approach to magnification dependent and independent classification of breast cancer in histopathological images. *Biomed. Signal Process. Control* **2021**, *63*, 102192. [\[CrossRef\]](#)
19. Negassi, M.; Suarez-Ibarrola, R.; Hein, S.; Miernik, A.; Reiterer, A. Application of artificial neural networks for automated analysis of cystoscopic images: A review of the status and prospects. *World J. Urol.* **2020**, *38*, 2349–2358. [\[CrossRef\]](#)
20. Das, R.; Sailo, L.; Verma, N.; Bharti, P.; Saikia, J.; Imtiwati; Kumar, R. Impact of heat stress on health and performance of dairy animals: A review. *Vet. World* **2016**, *9*, 260–268. [\[CrossRef\]](#)
21. Swati, Z.N.K.; Zhao, Q.; Kabir, M.; Ali, F.; Ali, Z.; Ahmed, S.; Lu, J. Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput. Med. Imaging Graph.* **2019**, *75*, 34–46. [\[CrossRef\]](#) [\[PubMed\]](#)

22. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009.
23. Yang, L.; Shami, A. On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice. *arXiv* **2022**, arXiv:2007.15745. [[CrossRef](#)]
24. Bressem, K.K.; Adams, L.C.; Erxleben, C.; Hamm, B.; Niehues, S.M.; Vahldiek, J.L. Comparing different deep learning architectures for classification of chest radiographs. *Sci. Rep.* **2020**, *10*, 13590. [[CrossRef](#)] [[PubMed](#)]
25. Pelikan, M.; Goldberg, D.E.; Cantú-Paz, E. BOA: The Bayesian optimization algorithm. In: Proceedings of the Genetic and Evolutionary Computation Conference. *Citeseer* **1999**, *1*, 525–532. [[CrossRef](#)]
26. Sameen, M.I.; Pradhan, B.; Lee, S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. *Catena* **2020**, *186*, 104249. [[CrossRef](#)]
27. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 3320–3328. [[CrossRef](#)]
28. Subramaniam, M.; Narasima Prasad, L.V.; Janakiramaiah, B.; Mohan Babu, A.; Sathishkumar, V.E. Hyperparameter Optimization for Transfer Learning of VGG16 for Disease Identification in Corn Leaves Using Bayesian Optimization. *Big Data* **2022**, *10*, 215–229. [[CrossRef](#)]
29. Siemon, M.S.N.; Shihavuddin, A.S.M.; Ravn-Haren, G. Sequential transfer learning based on hierarchical clustering for improved performance in deep learning-based food segmentation. *Sci. Rep.* **2021**, *11*, 813. [[CrossRef](#)]
30. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
31. Bezerra, L.A.; Ribeiro, R.R.; Lyra, P.R.M.; Lima, R.C.F. An empirical correlation to estimate thermal properties of the breast and of the breast nodule using thermographic images and optimization techniques. *Int. J. Heat Mass Transf.* **2020**, *149*, 119215. [[CrossRef](#)]
32. Gayathri, S.L.; Bhakat, M.; Mohanty, T.K.; Chaturvedi, K.K.; Kumar, R.R.; Gupta, A.; Kumar, S. Udder thermogram-based deep learning approach for mastitis detection in Murrah buffaloes. *Comput. Electron. Agric.* **2024**, *220*, 108906. [[CrossRef](#)]
33. Brito, J.R.F.; Caldeira, G.A.V.; Verneque, R.S.; Brito, M.A.V.P. Sensitivity and specificity of the California Mastitis Test as a diagnostic tool for subclinical mastitis in quarter somatic cell count estimation. *Braz. J. Vet. Res.* **1997**, *17*, 49–53. [[CrossRef](#)]
34. Macêdo, D.; Dreyer, P.; Ludermir, T.; Zanchettin, C. Training Aware Sigmoidal Optimizer. *arXiv* **2021**, arXiv:2102.08716v1.
35. de Vasconcelos, J.H.; dos Santos, W.P.; de Lima, R.C.F. Analysis of Methods of Classification of Breast Thermographic Images to Determine their Viability in the Early Breast Cancer Detection. *IEEE Lat. Am. Trans.* **2018**, *16*, 1631–1637. [[CrossRef](#)]
36. Abiwinanda, N.; Hanif, M.; Hesaputra, S.T.; Handayani, A.; Mengko, T.R. Brain Tumor Classification Using Convolutional Neural Network. In Proceedings of the World Congress on Medical Physics and Biomedical Engineering 2018, Prague, Czech Republic, 3–8 June 2018; Springer: Singapore, 2019; Volume 1, pp. 183–189. [[CrossRef](#)]
37. Mehrotra, R.; Ansari, M.A.; Agrawal, R.; Anand, R.S. A Transfer Learning approach for AI-based classification of brain tumors. *Mach. Learn. Appl.* **2020**, *2*, 100003. [[CrossRef](#)]
38. Martins, R.F.S.; Paim, T.; Dallago, S.L.B.; Melo, C.B.; Louvandini, H.; McManus, C. Mastitis detection in sheep by infrared thermography. *Res. Vet. Sci.* **2013**, *94*, 722–724. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.