On Farm Automatic Sheep Breed Classification using Deep Learning

Sandeep Mandapati
Department of EECS
Florida Atlantic University
Boca Raton, Florida, USA
smandapati2022@fau.edu
Z23684014

Subash Gupta Karamsetty
Department of EECS
Florida Atlantic University
Boca Raton, Florida, USA
skaramsetty2022@fau.edu
Z23689645

Akhil Yarlagadda
Department of EECS
Florida Atlantic University
Boca Raton, Florida, USA
ayarlagadda2022@fau.edu
Z23689300

***Abstract*—Sheep producers rely on accurate breed identification to assess the commercial value of their flocks. However, inexperienced farmers often struggle with this task. In this paper, we present our contributions to this field, which involve the development of a convolutional neural network model that is deployed in a sheep farm setting. A comprehensive database comprising 1680 high-resolution images of sheep, encompassing four distinct breeds, is considered and used. Each image in the dataset has been meticulously labeled by an expert with its corresponding breed. Leveraging deep learning and CNN techniques, we have trained a sheep breed classifier that achieves an impressive average accuracy of 93.75% with the VGG16 model and then enhanced it using the normalization and VGG19 models to get an accuracy of 97.68% and 93.87%, respectively. Our system demonstrates promising potential for accurate and efficient breed identification in sheep farming operations, thereby aiding producers in making informed decisions regarding flock management and commercial valuation.**

***Index Terms*—Sheep breed identification, Deep learning, Convolutional Neural Networks, Sheep farming, Breed classification, Image database, VGG-16, VGG-19.**

I. Introduction

Profit for sheep producers is intricately tied to the commercial value of their flocks and the costs associated with rearing the sheep. In Australia, where sheep farming is a significant industry, the commercial value of a sheep is primarily determined by its meat weight, commonly referred to as carcass weight. However, farmers face a challenge in accurately estimating the productivity of their flocks since carcass weight can only be determined post-slaughter at the abattoir. Consequently, farmers typically rely on live weight measurements, which include gut fill, wool weight, and bone frame. Notably, sheep often experience a drop in weight of around 2 kg between being weighed off pasture and again after an overnight fast.

The studies have developed prediction models for wool growth rates, with an average root mean square error ranging between 6 and 16 g/day. Additionally, the variation in meat yields among different sheep breeds poses a significant concern for producers. For instance, a study comparing the meat production of shorn sheep from twelve different breeds found substantial differences in meat yield, even among sheep of similar live weights. This variation, such as Merino cross sheep yielding 0.7 kg less meat than Southdown cross sheep of equivalent live weight, underscores the potential losses incurred by enterprises with large sheep populations.

Despite these challenges, there is a notable absence of automatic sheep breed identification models in the literature. Thus, this paper aims to address this gap by focusing specifically on the development of an automatic sheep breed identification system. Given the subtle differences in appearance among sheep breeds, computer vision (CV) and machine learning (ML) techniques present promising solutions. However, existing CV applications in agriculture often overlook sheep farming due to difficulties in segmenting individual sheep from uniform-colored flocks, predicting deformable body shapes, and extracting features amidst wool presence. Common classifiers in CV applications include support vector machines (SVM), convolutional neural networks (CNN), and clustered polynomial regression models.

In addition to the replication of the original model results, we made a few enhancements to the original model by adding normalization layers and data augmentation of images. We also created a new model using VGG19, which is taken as a transfer learning model, and obtained a maximum testing accuracy of 97.68%. Given the absence of an automatic sheep breed classification system in this paper, we have developed and tested multiple models to obtain the most accurate model to help the sheep breeders.

II. Materials and Methods

A. Data acquisition and preprocessing

The dataset we considered consists of the images captured using a GoPro Hero5 camera, which is configured to record at 24 frames per second (fps) with a resolution of 1920×1080 pixels. Various camera positions were tested to optimize recording quality. A ”good quality” video was defined as one that provided an unobstructed view of the face and front of the sheep, with only one sheep captured in each frame to avoid segmentation issues. A total of 160 sheep from four different breeds were selected: Merino (70), Suffolk (63), White-face Suffolk (16), and Poll Dorset (11), as identified by the grazier. To ensure a balanced dataset, an equal number of frames were captured for each breed, totaling 420 frames per breed.

The dataset captured a range of variations, including differing sheep postures, brightness levels due to varying weather conditions, moving shadows, and image interference.

The following steps were undertaken for data preprocessing:
1. **Frame Extraction and Filtering**: Individual frames were extracted from the recorded videos and filtered to remove frames with no sheep or with only partial sheep.
2. **Frame Sampling:** To ensure diversity, frames were sampled by selecting every fourth frame.
3. **Labeling**: Each frame was labeled according to the breed of the sheep.
4. **Preprocessing Methods**: We have two types of datasets defined for the model. They are the full sheep dataset and the sheep face dataset.
a. Full Sheep Dataset: A Minimal preprocessing was applied to retain various types of variation. Images were cropped to achieve a uniform aspect ratio of 1080×1080 pixels and resized to 224×224 pixels to match the input size of the CNN model. We have 1619 images in this dataset.
b. Sheep Face Dataset: The preprocessing aimed to reduce irrelevant features by cropping only the sheep's face, eliminate posture variation by aligning all face images, and re-scale images to a uniform size. This involved selecting control points (the sheep's right eye and septum), inferring a geometric transformation matrix, and applying translation, rotation, and scaling to align the images. Images were also resized to 224×224 pixels. We have 1680 images in this dataset.
5. **Data Splitting**: The image dataset was split into a training set (80%) and a testing set (20%). The (80%) training dataset is again split into (20%) validation dataset and the (80%) training dataset.
6. **Data Augmentation**: Augmentation techniques, including translations, were applied to the training set to increase dataset variability. This is applied to the model when we have made the augmentation changes to it. The width and height shifting are performed to augment the image dataset and increase its size.
To assess the quality of sheep face image alignment, a template image was chosen where the right eye and nostrils were aligned diagonally. Three different sheep images were aligned to this template, and one color channel (red, green, or blue) was extracted from each. These channels were superimposed to create a single RGB image, allowing evaluation of the alignment quality based on the resulting face image. 7. **System Used**: A Google Colaboratory with a GPU instance is used for running the model and obtaining the accuracy for each model.

B. Implementation of Transfer Learning and Fine-Tuning

Transfer learning was employed using the VGG-16 model to classify four breeds of sheep. Only the last six layers of VGG-16 were fine-tuned in this study to construct a sheep breed classifier. To adapt the classification segment of the network, the final three layers of the network were replaced by three new fully connected layers. Similarly, as an advancement to the existing base paper, we have also employed the VGG-19 model to classify the four breeds of sheep.

To achieve the maximum accuracy's and F1-Scores, we have customized the learning rates across different layers and manipulated the number of neurons in the output layers to achieve them. We have frozen the first ten layers and did not make any modifications to the learning rates to maintain the pre-trained model behavior. For the next few layers, i.e., the middle layers, we have modified the learning rates, and for the final three layers, we have modified the number of neurons and the learning rates to obtain optimal accuracy.

C. Evaluation metrics

The performance of the developed model was assessed using five-fold cross-validation, providing a statistical estimation of its real-world efficacy with new sheep images. Implementation of five-fold cross-validation involved partitioning the data into five equal subsets, with four utilized for training and one for testing in each iteration. This process was repeated five times, ensuring that each subset served as the testing set exactly once.

Evaluation of the model's performance relied on both average accuracy and standard deviation metrics. These metrics offer comprehensive insights into the model's effectiveness across different subsets of the data, accounting for variations in performance across multiple testing scenarios. The utilization of five-fold cross-validation enables robust assessment of the model's generalization capability and ensures reliable performance estimation in practical farm settings, where variability in sheep appearance and environmental conditions is inevitable.

III. Performance Evaluation and Analysis: Models, Experiments, and Metrics

In order to replicate the results of the base model from the paper and then enhance it to get better performances, we have created four different models.

A. Model 1: Fine tuning of the VGG-16

This model is a direct replication of the model discussed in the paper. We have considered both datasets: the full sheep dataset and the sheep face dataset. Initially, we have considered the sheep face dataset, taken a VGG-16 model using transfer learning, and set the include_top to false to remove the fully connected layers and the output layer. Then we have frozen the first 10 layers of the model, and the middle layers, i.e., the next 3 layers, were given a learning rate of 0.0001, and we have added two dense layers with 64 neurons each (to this conclusion after multiple trial and error options to replicate and get near to the paper's accuracy) and set a learning rate of 10 for these layers. We also passed the full sheep dataset as an input to the same model and obtained the accuracy and other performance parameters.

B. Model 2: Pre-trained VGG-16 with a SVM Classifier

This model closely follows the one outlined in the research paper. We've worked with two datasets: the full sheep dataset and the sheep face dataset. Starting with the sheep face dataset, we employed a VGG-16 model using transfer learning. Instead of training the entire VGG-16 model from scratch, we utilized its pre-trained version and extracted features specifically from its fc8 layer. These features were then fed into a Support Vector Machine (SVM) for classification purposes. Essentially, we transformed the VGG-16 model into a feature extractor, allowing the SVM to learn and classify based on these extracted features. This approach is particularly useful for leveraging the powerful feature extraction capabilities of deep learning models like VGG-16 while utilizing the flexibility and efficiency of SVMs for classification tasks. We also passed the full sheep dataset as an input to the same model and obtained the accuracy and other performance parameters.

C. Model 3: Fine tuning of the VGG-19

In the previous two models, we used the VGG-16. Now we have considered the VGG-19 model. Again, we have used both datasets, which are the full sheep dataset and the sheep face dataset. Initially, we have considered the sheep face dataset, taken a VGG-19 model using transfer learning, and set the include_top to false to remove the fully connected layers and the output layer. Then we have frozen the first 12 layers of the model. Now, the middle layers, i.e., the next 4 convolution layers, were given a learning rate of 0.001, and we have added two dense layers, one with 512 neurons and the other with 1024 neurons (to this conclusion after multiple trial and error options to replicate and get near to the paper presented accuracy), and set a learning rate of 0.001 for these layers. We also passed the full sheep dataset as an input to the same model without changing any parameters and obtained the accuracy and other performance parameters. The model diagram is shown in Figure 1.

D. Model 4: VGG-16 with Normalization and Data Augmentation

Fine tuning of the VGG-16 model is considered by removing the top layers, and in order to obtain better accuracy and lower loss percentages, we have added normalization and data augmentation along with the convolution layers. Similar to the previous model, we have frozen the first 10 layers of the model. Then, for the output of the model, we have added a series of convolutional layers followed by batch normalization and ReLU activation functions. After three consecutive convolutional layers with 512 filters each, 3x3 kernel size, and 'same' padding, max pooling is applied to down-sample the feature maps. Subsequently, a global average pooling layer is employed to reduce the spatial dimensions of the feature maps. Then the fully connected layers with two dense layers, each containing 64 neurons, followed by batch normalization, ReLU activation, and dropout regularization are added to prevent over-fitting and obtain the output. Moreover, to augment the dataset and boost the model's robustness, we implemented data augmentation techniques such as width and height shifting. By applying transformations to the training data, such as horizontally and vertically translating images by 10% of their total dimensions, we aimed to diversify the dataset and improve model generalization. This augmentation

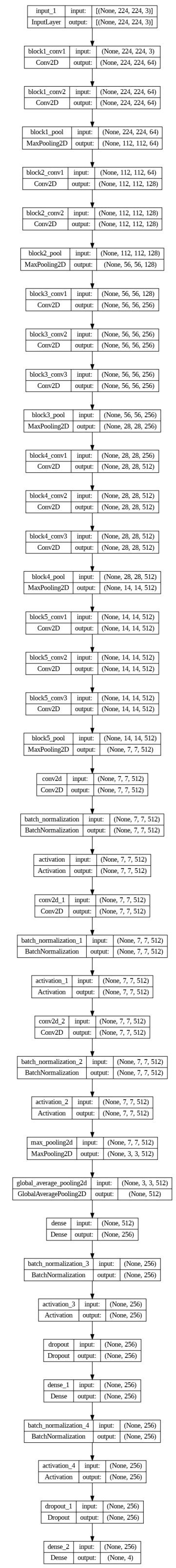

Fig. 1. Model 2: Pre-trained VGG-16 with a SVM Classifier

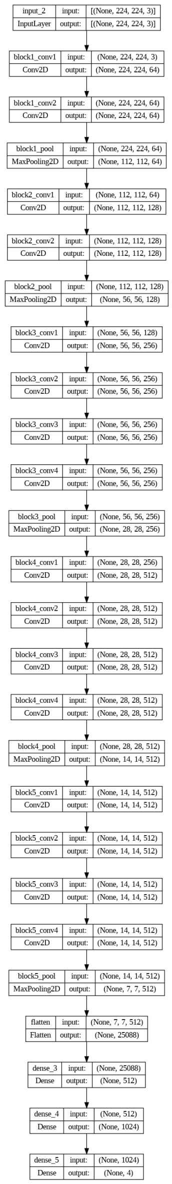

Fig. 2. Model 3: Fine tuning of the VGG-19

process contributes to the overall performance enhancement of the model. Here, the learning rate for the convolutional layers is 0.0001, and for the final layers, it is 1.0. The model diagram is shown in Figure 2.

The performance metrics obtained for different models when the two datasets are given as input and tabulated are shown in tables 1 and 2, with barplots 3 and 4.

IV. Results

The optimal camera position for farm monitoring was determined to be at the entrance of the weighing station within the drafting unit. This strategic placement capitalized on the behavior of sheep, which tended to lower their faces upon entering the weighing station. Additionally, this position facilitated multiple captures of individual sheep as they spent extended periods within the camera's view compared to other positions.

The impact of fine-tuning different numbers of layers in the VGG-16 and VGG-19 models was investigated to assess its influence on the performance of the sheep breed classifier. Analysis of accuracy curves revealed a gradual increase in both training and validation accuracy, converging to a high final percentage. Similarly, the loss curves demonstrated continuous

TABLE I
Full Sheep Image Dataset

Models	Performance Metrics		
	SD of Accuracy	*Accuracy*	*F1 Score*
Fine Tuning of VGG-16	0.0157	0.9074	0.9075
VGG-16 with SVM Classifier	0.0147	0.9549	0.9549
Fine Tuning of VGG-19	0.0089	0.9209	0.9210
VGG-16 With Normalization	0.0081	0.9537	0.9536

[a] All table values are the average values of the five folds.

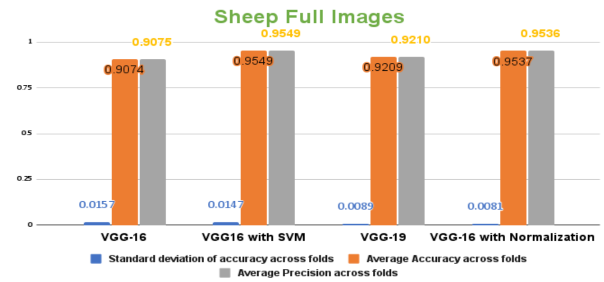

Fig. 3. A bar graph represents the performance metrics of various models when considering the Sheep Face dataset.

TABLE II
Sheep Face Image Dataset

Models	Performance Metrics		
	SD of Accuracy	*Accuracy*	*F1 Score*
Fine Tuning of VGG-16	0.0193	0.9375	0.9372
VGG-16 with SVM Classifier	0.0091	0.9685	0.9683
Fine Tuning of VGG-19	0.0061	0.9387	0.9383
VGG-16 With Normalization	0.0089	0.9768	0.9768

[a] All table values are the average values of the five folds.

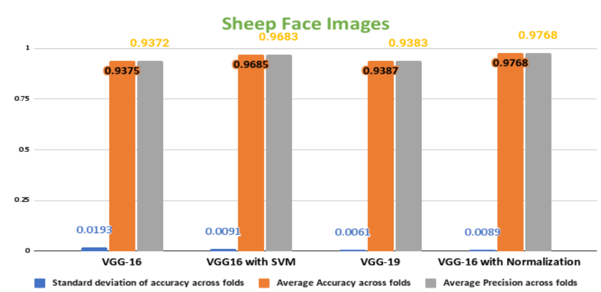

Fig. 4. A bar graph represents the performance metrics of various models when considering the Full Sheep dataset.

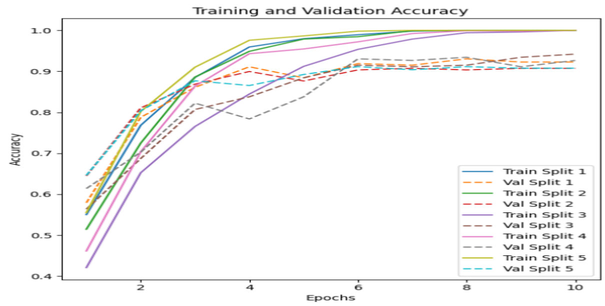

Fig. 5. The graph above represents the accuracy plot between training and validation for the Fine Tuning of VGG-16 model of the full sheep image dataset.

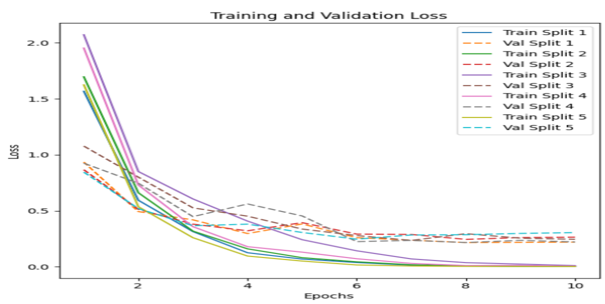

Fig. 6. The graph above represents the loss plot between training and validation for the Fine Tuning of VGG-16 model of the full sheep image dataset.

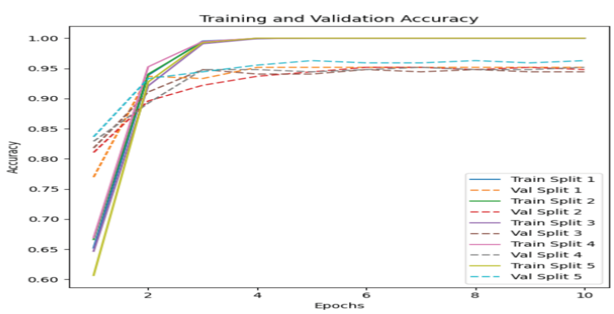

Fig. 7. The graph above represents the accuracy plot between training and validation for the Fine Tuning of VGG-16 model of the sheep face image dataset.

improvement after each iteration of optimization.

For Model 1, when we considered the full sheep dataset, we received an accuracy of 90.74%, which is specified in Table 1, and the plots for the training and validation accuracy are shown in Figure 5, and the training and validation loss are shown in Figure 6. Later, we considered the sheep face dataset and passed it to the same model without any parameter changes. In this case, we have received an accuracy of 93.75%, which is specified in Table 2, and the plots for the training and validation accuracy are shown in Figure 7, and the training and validation loss are shown in Figure 8.

For Model 2, when we considered the full sheep dataset, we received an accuracy of 95.49%, which is approximately 4.5% higher than the previous model, which is also specified

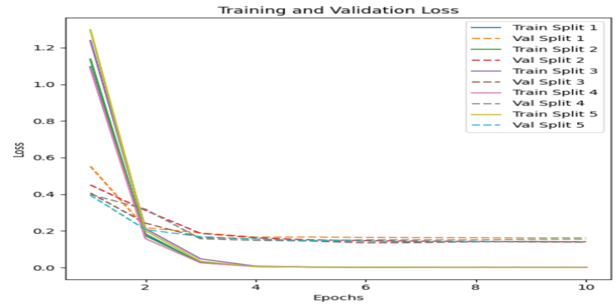

Fig. 8. The graph above represents the loss plot between training and validation for the Fine Tuning of VGG-16 model of the sheep face image dataset.

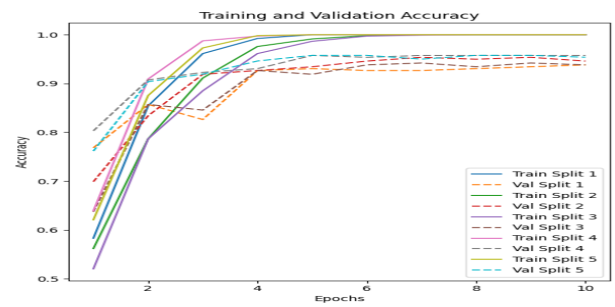

Fig. 9. The graph above represents the accuracy plot between training and validation for the Fine Tuning of VGG-19 model of the full sheep image dataset.

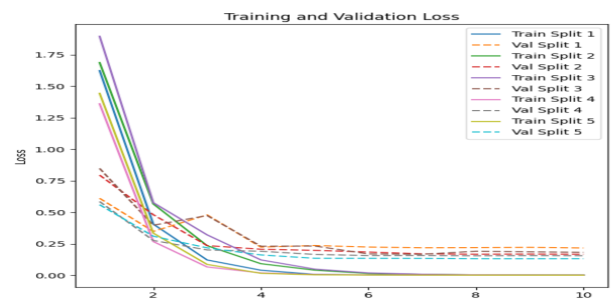

Fig. 10. The graph above represents the loss plot between training and validation for the Fine Tuning of VGG-19 model of the full sheep image dataset.

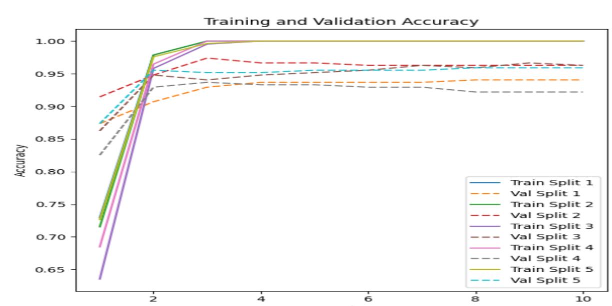

Fig. 11. The graph above represents the accuracy plot between training and validation for the Fine Tuning of VGG-19 model of the sheep face image dataset.

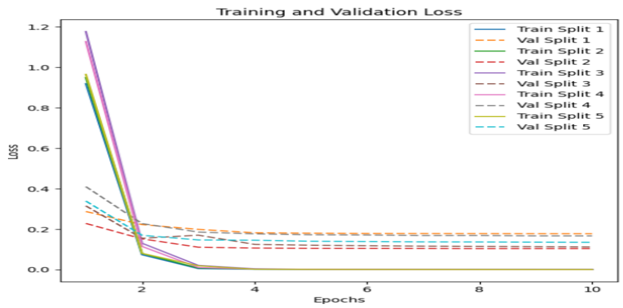

Fig. 12. The graph above represents the loss plot between training and validation for the Fine Tuning of VGG-19 model of the sheep face image dataset

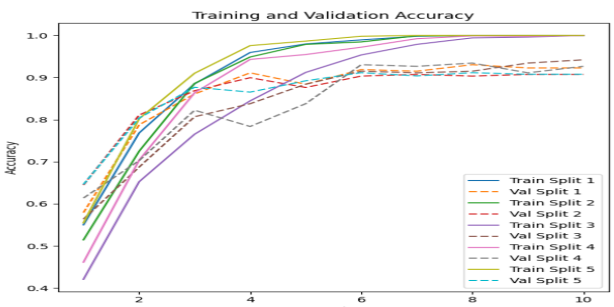

Fig. 13. The graph above represents the accuracy plot between training and validation for the VGG-16 model, which is enhanced with normalization and data augmentation of the full sheep image dataset.

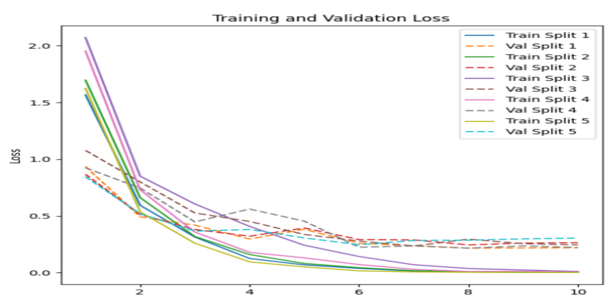

Fig. 14. The graph above represents the loss plot between training and validation for the VGG-16 model, which is enhanced with normalization and data augmentation of the full sheep image dataset.

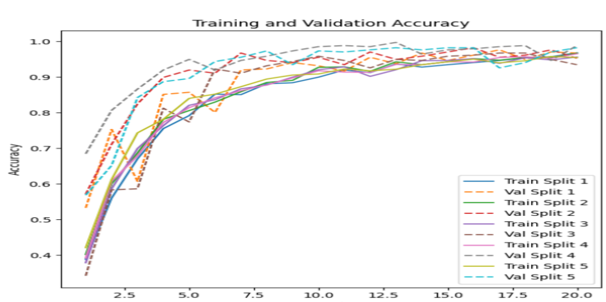

Fig. 15. The graph above represents the accuracy plot between training and validation for the VGG-16 model, which is enhanced with normalization and data augmentation of the sheep face image dataset.

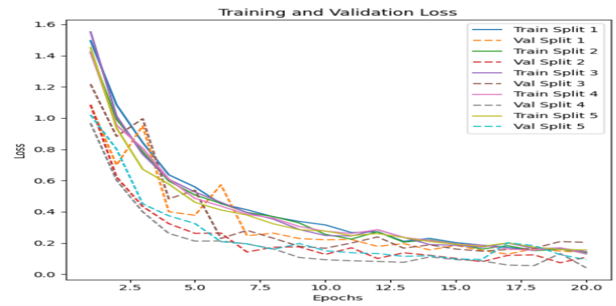

Fig. 16. The graph above represents the loss plot between training and validation for the VGG-16 model, which is enhanced with normalization and data augmentation of the sheep face image dataset.

in Table 1. Later, we considered the sheep face dataset and passed it to the same model without any parameter changes. In this case, we have received an accuracy of 96.85%, which is approximately 3% higher than the previous model, which is specified in Table 2.

For Model 3, when we considered the full sheep dataset, we received an accuracy of 92.09%, which is more than the model 1 which is also specified in Table 1, and the plots for the training and validation accuracy are shown in Figure 9, and the training and validation loss are shown in Figure 10. Later, we considered the sheep face dataset and passed it to the same model without any parameter changes. In this case, we have received an accuracy of 93.87%, which is not a major increase when compared with model 1, which is also specified in Table 2, and the plots for the training and validation accuracy are shown in Figure 11, and the training and validation loss are shown in Figure 12.

For Model 4, when we considered the full sheep dataset, we received an accuracy of 95.37%, which is specified in Table 1, and the plots for the training and validation accuracy are shown in Figure 13, and the training and validation loss are shown in Figure 14. Later, we considered the sheep face dataset and passed it to the same model without any parameter changes. In this case, we have received an accuracy of 97.68%, which is higher of all the models and which is also specified in Table 2, and the plots for the training and validation accuracy are shown in Figure 15, and the training and validation loss are shown in Figure 16.

V. Conclusion

In this study, various training parameters were analyzed, and experiments were conducted using two types of training images: full sheep images containing noise and significant variation, and cropped facial images with minimal noise and variation.

Our findings indicate that fine-tuning the last six layers of VGG-16 for 10 epochs resulted in a classification accuracy of 93.75%, with a standard deviation of 1.93. Fine-tuning the last 7 layers of the VGG-19 for the same 10 epochs has resulted in a slightly higher classification accuracy of 93.87%. But fine-tuning the VGG-16 by adding a few additional convolution layers and performing batch normalization and data augmentation has helped in improving the accuracy by 3.93%, making it 97.68% the highest accuracy. This demonstrates the effectiveness and practicality of the model for on-farm use.

VI. References

[1] Armstrong, J.S., 2012. Illusions in regression analysis. *International Journal of Forecasting*, 28(3), 689–694.

[2] Asamoah Boaheng, M., Sam, E., 2016. Morphological characterization of breeds of sheep: a discriminant analysis approach. *SpringerPlus*, 5(1), 1–12.

[3] Atanbori, J., Duan, W., Murray, J., Appiah, K., Dickinson, P., 2016. Automatic classification of flying bird species using computer vision techniques. *Pattern Recognition Letters*, 81(C), 53–62.

[4] Bayramoglu, N., Heikkilä, J., 2016. Transfer learning for cell nuclei classification in histopathology images. In: Hua, G., Jégou, H. (Eds.), *Computer Vision – ECCV 2016 Workshops*. Springer International Publishing, Cham, pp. 532–539.

[5] Bunbury, 2018. WA: daily weather observations 2018. *Bom.gov.au*.

[6] Burke, J., Nuthall, P., McKinnon, A., 2004. An analysis of the feasibility of using image processing to estimate the live weight of sheep.

[7] Carneiro, H., Louvandini, H., Paiva, S., Macedo, F., Mernies, B., McManus, C., 2010. Morphological characterization of sheep breeds in Brazil, Uruguay and Colombia. *Small Ruminant Research*, 94(1), 58–65.

[8] Deng, J., Dong, W., Socher, R., Li, L., Li, K., Fei-Fei, L., June 2009. ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255.

[9] Devikar, P., 2018. Transfer learning for image classification of various dog breeds. *International Journal of Advanced Research in Computer Engineering & Technology*, 5(12), 2707–2715.

[10] Finlayson, J., Cacho, O., Bywater, A., 1995. A simulation model of grazing sheep: Animal growth and intake. *Agricultural Systems*, 48(1), 1–25.

[11] Gopalakrishnan, K., Khaitan, S.K., Choudhary, A., Agrawal, A., 2017. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 157, 322–330.

[12] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

[13] Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554.

[14] Hong, F., Tan, J., Mccall, D., 2000. Application of neural network and time series techniques in wool growth modeling. *Transactions of the ASAE*, 43(1), 139–144.

[15] Liu, J., Kanazawa, A., Jacobs, D., Belhumeur, P., 2012. Dog breed classification using part localization. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (Eds.), *Computer Vision – ECCV 2012*. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 172–185.