Small Ruminant Research 213 (2022) 106724

Contents lists available at ScienceDirect

Small Ruminant Research

journal homepage: www.elsevier.com/locate/smallrumres

Application of artificial intelligence and digital images analysis to automatically determine the percentage of fiber medullation in alpaca fleece samples

Max Quispe Bonilla [a,b,*], Luis Serrano-Arriezu [a,c], Jesús Daniel Trigo [a,c], Christian Quispe Bonilla [d], Adolfo Poma Gutiérrez [e], Edgar Quispe Peña [e,f]

[a] Institute of Smart Cities (ISC), Department of Electrical, Electronic and Communication Engineering, Public University of Navarre (UPNA), Arrosadia Campus, 31006 Pamplona, Spain
[b] Laboratory of Innovations and Technological Development, MAXCORP Technologies S.A.C, 866 Ruiseñores Av., Office 403. Santa Anita, Lima, Peru
[c] Navarra Institute for Health Research (IdiSNA), 31008 Pamplona, Spain
[d] Neural X S.A.C, 155 Kuelap Psg., La Molina, Lima, Peru
[e] Textile Fiber Laboratory, Natural Fiber's Tech S.A.C, 207Malaga Street., La Molina, Lima, Peru
[f] Department of Animal Production. Universidad Nacional Agraria La Molina, La Molina Av., La Molina, Lima, Peru

ARTICLE INFO

Keywords:
Deep learning
Medulla
Objectable fiber
South American camelids

ABSTRACT

The aim of this research is to develop and validate two computer programs based on artificial intelligence (AI) and digital image analysis (DIA) in order to determine the incidence of medullation in white alpaca fibers. Two data sets were analyzed: 76 samples of Huacaya alpaca fibers obtained from Huancavelica, Peru, and 200 samples of white alpacas of two genotypes (Huacaya, n =100; Suri, n = 100), obtained from Arequipa, Peru. The preparation of each sample followed the procedure described in IWTO-8-2011. The Pytorch framework was used to generate several training models based on the You Only Look at Once (YOLO) architecture. Circa 4000 pictures of fibers were taken and 661 of them were selected as representative. Using the LabelImg software, the fibers present in each representative picture (approximately 10 fibers/picture) were labeled as one of these two classes: either medullated or non-medullated. Subsequently, the data augmentation technique was applied to expand the data set to 3966 photographs. Thus, 90 of them were used as initial validation data, while the reaming 3876 pictures (containing a total of 23,964 labeled fibers) were used as training data. Matlab was used to develop the DIA-based software. More specifically, algorithms of pre-processing, segmentation, smoothing, skeletonization and Hough transform were implemented to detect medullated and non-medullated fibers. Correlation and linear regression analyses were used to evaluate the models. The medullation percentage results show that there is no statistically significant difference between the AI-based method and the projection microscope method (p-value = 0.668 and 0.672 for the t-student and Wilcoxon tests, respectively). Moreover, the correlation of each of the developed computer methods with the projection microscope method is very strong (r = 0.99 and 0.97). This confirms the software ability to perform the recognition of fibers with and without medullation. Similar results (p-value = 0.357) were obtained when comparing the projection microscope method and DIA-based software method. Finally, using the proposed framework, the average time required to analyze a sample was 19.44 s. As a result, this software allows the implementation of practical, precise, and efficient methodologies to determine the incidence of medullation of alpaca fibers.

1. Introduction

The fleeces of South American camelids are composed of fibers both with and without the presence of medullation (Villarroel, 1963; Martinez et al., 1997). Among the latter stand out the fibers of continuous medullation and the strongly medullated, also known as objectionable

* Corresponding author at: Laboratory of Innovations and Technological Development, MAXCORP Technologies S.A.C, 866 Ruiseñores Av., Office 403. Santa Anita, Lima, Peru.

E-mail address: maxdavid22@gmail.com (M.Q. Bonilla).

https://doi.org/10.1016/j.smallrumres.2022.106724
Received 27 December 2021; Received in revised form 23 March 2022; Accepted 11 May 2022
Available online 14 May 2022
0921-4488/

fibers (Hunter et al., 2013). The presence of these type of fibers is often undesirable for the textile industry due to a number of reasons. First, the wide medulla of these fibers makes them dye with a shade paler than those non-medullated. This effect is enhanced by the reflection of light from the surface of the medulla, which makes undyed fibers appear chalky white (Shelton, 1995; Balasingam, 2005). Moreover, in contact with the skin, their coarseness causes an unpleasant sensation to people, commonly referred to as "prickle". This is due to the thickness of some fibers (those bigger than 30 μm) protruding from (Frank et al., 2014). In addition, they give a heterogeneous appearance in yarns and fabrics (Wang et al., 2005) that decreases the price of the product (McGregor, 1997). This is so because they break quickly during the processing of the textiles, increasing thereby the proportion of waste (Gupta et al., 1981). The prickle and heterogeneity of fabrics are more intense as the incidence of medullation increases. However, medullation is not always a defect, given that its presence provides a high insulating and thermal quality to the garments (Wang et al., 2005). Moreover, it makes them lighter, which is why fabrics based on alpaca fibers are lighter than those of wool. This is explained by the differences in density, being the former around 1309 g / cm^3 and the latter 1320 g / cm^3 (Czaplicki, 2012).

Thus, it is of utmost importance to determine the incidence of medullated fibers. To do so, the projection microscope (PM) procedure is the standard, direct and objective method. However, it is laborious, expensive, time consuming (Lupton and Pfeiffer, 1998; Balasingam, 2005). Shakyawar et al. (2013) found that the average time taken by two operators to measure 500 fiber sites with PM is apparoximately two hours and the results may vary from person to person. Therefore, it does not have an extensive and practical use (Shakyawar et al., 2013). Efforts aimed at measuring the degree of medullation can be found in the literature. The most relevant pieces of equipment and procedures include: (a) the Wronz Medulometer, based on the photoelectric technique (Lappage and Bedford, 1983; Wood, 2003; Balasingam, 2005); (b) modifications of the Optical Fibre Diameter Analyser (OFDA), which incorporate lighting and software based on opacity (Brims and Peterson, 1994; Turpie and Steenkamp, 1995); (c) the near-infrared ray reflectance analysis (NIRS), based on spectrometry (Boguslavsky at al., 1992); (d) the sonic digitizer technique (Blakeman et al., 1988), and (e) the automatic imagen analysis system (Qi et al., 1995).

However, these equipments have been developed mainly for the evaluation of wool and mohair fibers (Lappage and Bedford, 1983; Blakeman et al., 1988; Boguslavsky et al., 1992; Brims and Peterson, 1994; Qi et al., 1995; Balasingam, 2005; Cottle and Baxter, 2015). Little evaluation has been performed for cashmere and no validation tests have been conducted on South American camelid fibers. To date, the OFDA100 is the only device available that, in compliance to IWTO-57-2000 (IWTO, 2017b), evaluates the incidence of medullation of wool fibers in a practical, fast and accurate way. However, according to the different authors, the results in mohair fibers are still controversial (Lupton and Pfeiffer, 1998; Botha and Hunter, 2010; Cottle and Baxter, 2015). In addition, in alpacas fibers, it has been shown to have unadequate accuracy and precision (Pinares et al., 2018; Torres, 2020). Moreover, all fibers of Angora rabbit have at least one medulla canal, however Rafat et al. (2007) found mean percentage of medullated fibers very low, ranging from 0.1% to 7.3%, therefore they concluded that a new definition about opacity must be developed to mesure medullation in Angora rabbit fiber.

Along with the growth of computer science and digital image analysis, there has been research efforts towards the development of algorithms to measure the diameter of animal fibers, as well as some other variables (Baxter et al., 1992; Qi et al., 1995; Deng and Ke, 2010; Arcidiácono et al., 2014; Quispe et al., 2017). In the topic of medullation, Qi et al. (1995) and Shakyawar et al. (2013) used a PM together with an image analysis software to determine, with precision and accuracy, the percentage of heterotopic fibers, thick and kemp type, in the sheep and goat fleeces, and in wool of creole sheep of India. However, there are still no reports of any device, method or procedure created to specifically evaluate the medullation of alpaca fibers.

At present, Artificial Intelligence (AI) has shown to be efficient for the recognition of arbitrary shaped objects. AI requires a relatively large amount of data (usually referred to as "training set") or, at least, enough input information to apply data augmentation techniques, and thus avoid overfitting in training (Krizhenysky and Sutskever, 2012). Therefore, this technology could be useful for the recognition of medullated and non-medullated fibers. Nevertheless, no previous work in the literature has approached the fiber medullation issue by means of AI.

Therefore, there is a lack of devices and procedures that would allow the practical identification and swift counting of medullated fibers in alpacas and llamas. Producers are unable to breed animals with good quality fleeces because they do not have enough objective information of the animals fibers. The textile industry must draw upon additional processes (such as dehairing or other treatments) to remove strongly medullated or objectionable fibers, which in many cases are inefficient (Hack et al., 1999; Frank et al., 2014). Dehairing is a procedure that works around the problem in llamas and guanacos, but it is arduous, expensive, unhealthy and has variable results (Tarqui, 2008; Quispe et al., 2015). However, in alpacas and llamas, it does not have good results (Wang et al., 2003). Therefore, the selection in favor of reducing the percentage of strongly medullated fibers is the best solution strategy (Gupta et al., 1981; Pinares et al., 2019; Cruz et al., 2018; Torres, 2020; Radzick-Rant and Wierckinska, 2021).

With all the considerations above, this work is aimed at developing two computer programs based on DIA and AI to determine the incidence of medullation in Huacaya and Suri alpaca fibers and evaluate their performance in terms of accuracy, precisión and time required compared to existing methodologies.

2. Materials and methods

2.1. Location and time frame of the research

The design and development of the computer programs was carried out in the Laboratory of Innovations and Technological Development of Maxcorp Technologies S.A.C (Lima, Peru), while the analysis of the fiber samples using the PM was carried out in the Textile Fiber Laboratory of Natural Fiber's Tech S.A.C (Lima, Peru). The fiber samples were obtained from two production units located in the regions of Huancavelica and Arequipa (Peru). The research was carried out between January to September, 2020.

2.2. Samples preparation

Two data sets of fibers were used in this research. First, 76 samples of Huacaya alpaca fibers from the Lachocc Research and Production Center (National University of Huancavelica, Peru) were analyzed. More specifically, 40 white and 36 light beige fibers were considered. These colors were chosen since they allow observing the medulla without the need of discoloring the fiber. Then, 200 samples of white alpacas of two genotypes (Huacaya, n = 100; Suri, n = 100), obtained from the Pacomarca Genetic Center belonging to INCATOPS, located in Arequipa, Peru, were also evaluated. It is worth noting that while Huacaya alpacas produce crimpy, curly, dense and soft wool, Suri alpacas have longer cylindrical locks resembling dreadlocks. All samples were taken from the mid-side area, which is located over the third last rib, half-way between the mid-line of the belly and the mid-line of the back at the height of the tenth dorsal rib (Aylan-Parker and McGregor, 2002; McGregor et al., 2011). The fibers had an average diameter of 21.72 ± 0.21 μm (minimum: 15.67 μm; maximum: 28.39 μm).

The preparation of each sample to be analyzed followed the procedure described in IWTO-8-2011 (IWTO, 2017a). The samples were washed in a solution of 7 parts 96% ethyl alcohol and 3 parts benzene. Then, they were dried with a towel by pressing with a roller. Fiber fragments were obtained from each sample whose lengths varied between 0.4 and 0.8 mm for which the Hardy micrometer was used. The fiber fragments were placed on a slide and dispersed, using a stirring rod and immersion oil. Then, a coverslip was placed on top. Care was taken that each sample image contained between 5 and 20 fragments of fibers, avoiding as much as possible fiber crosses and/or agglomeration of fiber fragments.

2.3. Model development using AI

Several training models were generated based on the You Only Look at Once (YOLO) architecture (Ultralytics, 2020) to develop the AI-based software. In the approach developed, the configuration was carried out using Python as programming language and PyTorch, an open-source machine-learning framework.

The image database stems from 40 fiber samples. Approximately, 100 photographs were taken from each fiber sample and each photograph contained an average of 5–10 fibers. Thus, 4000 photographs, with about 40,000 fibers altogether, were available. The photographs were obtained with a PM, equipped with a 4X magnifying lens. Out of the 4000 photographs, only 661 representative photographs with different percentages of medullated fibers (hereinafter, %MedFib) were taken in consideration for balancing data. Then, each one of the fibers was manually labeled as one of these two classes: either medullated fibers (MedFib) or non-medullated fibers (UnMedFib). Using the LabelImg software in YOLO format as a tool, the labeling consisted of graphically enclosing each fiber in a bounding box and indicating whether it was "medullated" or "not medullated".

Subsequently, the data augmentation technique was applied to the photographs with labeled fibers. This was carried out using the "Albumentations" library, part of the PyTorch ecosystem. The process consisted in creating new images by making random modifications in color, lighting, blur and rotation. Thus, the database is increased and the accuracy at inferencing is improved, making it more robust to the indicated changes. By this method, it was possible to obtain a total dataset of 3966 images, composed of the initial images (661) and those obtained as a result of the modifications (3305). Out of the 3966 images, 90 of them were used as data for the initial validation, while the reaming 3876 pictures (containing 23,964 labeled fibers) were used as training data for the models.

For the training process, the “Colaboratory” tool from Google was used. This is a powerful, web-based virtual machine with 25 GB of Video Random Access Memory (VRAM) and V100 or P100 graphics processing units (GPUs) available. Thus, from several model developed only two models (YOLOv5s and YOLOv5l) were chosen. These varied in the number of parameters assigned to the neural network (7.3 and 47.0 million) and sizes (14.13 and 90.87 MB). Then, the trained models were evaluated, by comparing the percentages of medullation obtained with the results of the PM procedure. For this purpose, only 40 samples of white alpaca fibers were used. For the model evaluation process, a laptop with an NVIDIA 1660 Ti GPU with 8 GB of VRAM and a 9th generation Core I7 with 16 GB of RAM was used. The machine used is relevant since, during this process, the average time used to evaluate a sample was measured.

2.4. Model development using DIA

Initially, 1000 random images from a database of 7600 photographs, belonging to 76 samples of alpaca fiber, were selected. A number of digital image operations were performed over the pictures, in a sequence as follows. First, they were transformed to grayscale to distinguish the poorly lit fibers. Second, they were improved by preprocessing-enhancement methods. Third, segmentation and smoothing operations were performed. Fourth, the edges of the fibers were detected and the respective filling with pixels of values equal to one (“1”) was performed. Fifth, since unwanted remains were present in the background of the images, such as grease and dirt, morphological operations of erosion and dilation were carried out to eliminate them and homogenize the images of the fibers. Sixth, the fibers were skeletonized, a process that consists in removing from a pattern as many pixels as possible without affecting the general shape thereof. In our approach, the objective was to obtain a skeleton (i.e. a line) with a thickness of a single pixel, connected and centered in the middle of the fiber. However, when applying this process, protruding branches of the skeleton were found. Therefore, the seventh step consisted in applying the process called pruning, which transformed the sample of the fiber into curves of 1-pixel thick. Eighth, the Hough Transform was used in order to find straight lines along the samples of the fiber images. Ninth, these lines were labeled and, finally, the counting of objects (in our case, fibers) began, resulting in the total number of fibers (hereafter, TFib).

To measure the diameter, the following steps were performed. Once the lines of the fibers were found, the algorithm checked if they were vertically, horizontally or obliquely disposed. Then, a sweep was performed; perpendicular to both ends and along these lines until finding a zero (“0”). Thus, the sum of these points (pixels) is the total diameter of the fiber.

To establish the number of medullated fibers (hereinafter, NºMedFib), the following procedure was applied. It relies on the fact that the interior of the fibers is transparent to visible light due to the immersion oil, given that it has the same refraction index. In addition, some fibers have a medulla, which is opaque and can be distinguished because of having holes or lines of dark color present inside the fiber after the original image is binarized (0 = medulla and edges, 1 = body of the translucent fiber and background). Thus, we proceeded to superimpose the lines previously found over the original binarized and inverted image. Using the algorithms previously developed, a certain diameter is measured in case the fiber has a medulla. Otherwise, this value is equal to "0". Then, the amount of measurements of diameters different that "0" is counted, obtaining thereby the NºMedFib. The percentage of medullated fibers (from now on, %MedFib) is the ratio between the NºMedFib and the TFib found, expressed as a percentage ($\%MedFib = 100 * N^{\circ}MedFib \ / \ TFib$).

All these programming codes, as well as the graphical user interface were developed in MATLAB programming language, under the Windows 10 operating system, with a Core i7 processor computer and 12 GB RAM. The machine used is relevant since, analogously to the AI-based model case, the average time per sample used to perform the evaluation was measured.

2.5. Validation of computer programs

Initially, 76 samples of alpaca fibers from Huancavelica were evaluated using a PM. 100 images of each sample were captured, thereby accumulating 7600 images of alpaca fiber samples. The validation process consisted in comparing two alternative methods to classify the fibers in the pictures as MedFib or UnMedFib and, hence, obtain the percentages of medullation of each sample. On one hand, a visual and manual recognition was performed, following the procedure in test method IWTO-8-2011 (IWTO, 2017a), but slightly modified (since the fibers were classified as medullated and non-medullated, only). On the other hand, the same images were evaluated using the developed computer programs, one based on AI and the other based on DIA, obtaining the percentages of medullation of each sample, by type of procedure used.

Subsequently, 200 samples of Huacaya and Suri alpacas were used. Each sample was divided into two parts. One of the subsamples was evaluated with AI-based software and the samples were prepared according IWTO-8-2011 (IWTO, 2017a). The other 200 subsamples were evaluated with an OFDA100 equipment, following the procedure described in IWTO-57-2000 (IWTO, 2017b).

2.6. Statistical analysis

Correlation and linear regression analyses were used to evaluate the models. Before, the %MedFib, obtained using the developed software (based on both AI and DIA), the PM and the OFDA100, were initially subjected to tests of normality and homogeneity of variances using Komogorov-Smirnov and Levene tests, respectively. Subsequently, different statistical methodologies were used to obtain a number of %MedFib-related descriptive statistics. To compare and evaluate the degree of relationship between the automatic computational methods and those obtained with the PM, the paired Student's t-tests and the Wilcoxon test were applied; simple linear regression analysis and Pearson's correlation (r) analysis were also applied, complemented by their corresponding scatter plots. To evaluate the existence of biases between computational methods and PM, functional regressions of geometric mean (FRGM) and simple linear regressions (SLR) were performed, of the differences and the average of the %MedFib obtained with PM and each of the computational methods, following the methodology indicated in IWTO-0-2012 Appendix B (IWTO, 2017c), complemented with their corresponding scatter plots. The %MedFib of the 200 samples obtained with AI-based software and OFDA100 were compared with t-test by genotypes (Huacaya and Suri), because the characteristics of their fibers differ. Likewise, the statistical analysis was complemented by a scatter plot, lineal regression and, Spearman's correlation. To perform these analyses, the software for statistical computing and graphics R was used (R Core Team, 2016).

3. Results

3.1. AI-based software

After evaluating about 1000 fibers per sample, the YOLOv5s model achieved an average processing speed of 8.82 s/sample, while the YOLOv5l model was slower (19.44 s/sample). As regards to the ability of measuring the %MedFib, both models showed a similar Pearson correlation (0.999 and 0.994) when comparing them to the PM procedure. For the regression analysis, the slope and intercept were respectively 0.967 and 3.568 for YOLOv5s and 0.991 and 0.976 for YOLOv5l. Based on these results, the YOLOv5s model was used for the development of the respective software, due to its speed, similarity of slope and high correlation.

The software developed calculates the TFib per sample and the number of medullated (N°MedFib) and non-medullated (N°UnMedFib) fibers, as well as the respective percentages, with an automatic, fast and practical process. However, the final processing speed is highly dependent on the computer used, being to a large extent determined by the computation power of the video card. For a fast processing, it should contain Compute Unified Device Architecture (CUDA) cores.

A benefit of the software developed is its ability to identify the types of fibers (medullated and non-medullated) by marking them in a box in the processed images, which can be saved according to the need of the user, thus allowing to verify the identification process.

3.2. DIA-based software

This software leverages digital image processing to analyze pictures of fiber samples. The captured images are processed one after the other to automatically summarize the results in the %MedFib, TFib evaluated, N°MedFib and N°UnMedFib. In addition, it saves the fields of valid photographs for further processing. Installed on a computer with a 2.8 GHz Core I7 processor (4 CPUs) and 12 GB of RAM, the developed software takes 240 s/sample.

On the other hand, all the fiber information displayed in the graphical interface is automatically saved in an Excel sheet.

Table 1
Overall statistics of the percentage of medullation and number of fibers evaluated, obtained with the Projection Microscope (PM), the Artificial Intelligence (AI)-based software and the Digital Image Analysis (DIA)-based software, in 76 samples of Huacaya alpaca fiber.

Variables evaluated and evaluation methods	Mean	Standard error	Minimum	Maximum
Medullation Percentage				
●PM	59.26	2.98	6.43	98.94
●AI-based software	61.04	2.90	9.00	99.00
●DIA-based software	54.80	3.49	16.24	93.64
N° fibers evaluated/equipment				
●PM	763	20	379	1161
●AI-based software	996	27	402	1649
●DIA-based software	732	21	472	958

Table 2
Statistics of the paired parametric t-Student's test and Wilcoxon non-parametric test, among the percentages of medullation obtained with the computer programs based on Artificial Intelligence (AI) and Digital Image Analysis (DIA), compared with the methodology of the projection microscope (PM), in 76 samples of Huacaya alpaca fiber.

Statistics				t-Student's test	Wilcoxon test
●Comparison between the AI-Based computer software and the PM method					
	PM	AI	Mean	Difference	Difference
Mean (%)	59.26	61.04	60.15	-1.78	-1.54
– StandardDeviation (%)	25.99	25.25	25.58	2.70	
– Standard Error	2.98	2.90	2.93	0.31	
– Difference interval				[−9.99 6.43]	[−10.04 6.76]
– t-value/w-valor of the difference				-0.429	2772.50
– P-value				0.668	0.672
– Significance				NS	NS
●Comparison between the DIA-Based computer software and the PM method					
				t-Student's test	Wilcoxon test
	PM	DIA	Mean	Difference	
– Mean (%)	59.26	54.80	55.72	1.85	5.51
– StandardDeviation (%)	25.99	22.05	24.5	6.31	
– Standard Error	2.98	1.37	3.87	1.00	
– Difference interval				[−5.10 14.02]	[−4.88 15.35]
– t-value/w-value of the difference				0.925	1709
– P-value				0.357	0.274
– Significance				NS	NS

3.3. Validation

The results of evaluating the two computer programs (based on AI and DIA) compared to the PM method are illustrated in Table 1. The algorithms produce an average of the %MedFib of alpaca very close to that obtained with the PM (61.04% and 54.80%, respectively, compared to 59.26%). In addition, the TFib evaluated by AI-based software is greater than that of the other two methods (996 versus 732 and 763), and its standard error is lower compared with the DIA-based method and PM. In addition, between 15 and 18 samples were evaluated (average=18.57 min/sample) per day (8-hour work day) with the PM method, but with the DIA and AI-based software, the processes were 4.6 and 126 times faster, respectively.

As shown in Table 2, regarding the results of calculating the %MedFib, no statistical difference was found when comparing the AI-based software with the PM methods (p-value = 0.668 and 0.672 for the t-Student's and Wilcoxon tests, respectively.). Additionally, when evaluating the relationship between the values obtained with the procedures, a correlation and regression coefficient close to one (0.99 in

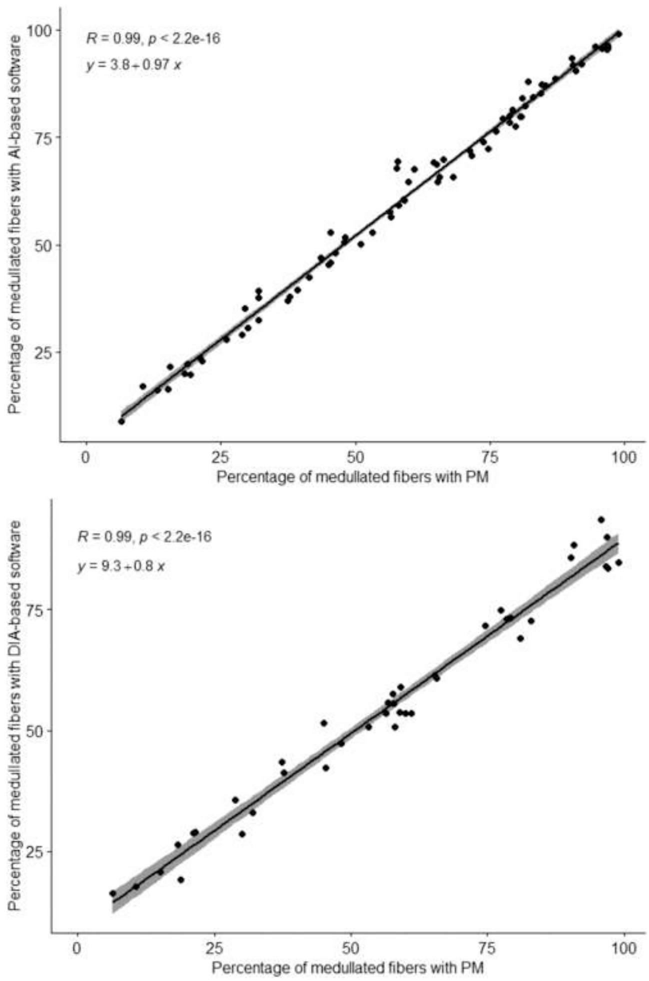

Fig. 1. Scatter plot of the percentages of medullated fibers (%MedFib) obtained with the AI- and the DIA-based software with the projection microscope (PM) methodology (top and bottom, respectively). Pearson's correlation coefficients and their significance are also shown, as well as the fit-line equation (AI= Artificial Intelligence; DIA: Digital Image Analysis).

both cases) was obtained, with an intercept not too far away of zero (−0.97), as can be seen in Fig. 1. Similar results were obtained when comparing the PM and DIA-based software methods. More specifically, the p-value is 0.36 and the Pearson's correlation index is 0.99 (Table 2 and Fig. 1).

The relationship between the differences and the averages of the % MedFib, considering the procedures using the developed software (based on both AI and DIA) and the PM method are shown in Table 3 and Fig. 2. The results show that, as the %MedFib increases, a slight increase in the differences between methods is also present. The increases in the differences are more noticeable in the DIA-based software than the AI-based software, since the regression coefficients (b_{xy}) compared to the PM method are 0.21 and 0.03, respectively. In addition, the correlation coefficients threw by the geometric functional analysis and the simple linear analysis are different from zero. The relationship degree is less manifest when the PM and the AI-based software are assessed, than when the PM and the DIA-based software are assessed (r = 0.27 and 0.80, respectively). Therefore, the bias lies between 7 and 3 times less with AI-PM than DIA-PM (when b_{yx} and r are considered for comparison, respectively).(Fig. 3).

Finally, the results of the evaluation of samples of white Huacaya and Suri alpaca fibers are shown in Table 4. It can be observed that the % MedFib| measured with both methods evaluated (AI-based software vs OFDA100) in Huacaya alpaca fibers differ greatly (54.96% and 10.94%, respectively, p-value <0.001). Similar results between methods were found in Suri alpaca fibers (39.86% and 8.69%, respectively, p-value < 0.001). Additionally, when evaluating the relationship between the results obtained with the OFDA100 and those obtained by the AI-based software, it was found that there is a high linear correlation (0.74), although this relationship is adjusted to a quadratic relationship, obtaining in this case a coefficient of determination of 0.54.

Table 3
Statistics of the Functional Regression of Geometric Mean (FRMG) and simple linear regression (SLR) of the differences vs averages percentage of medullation of Huacaya alpaca fibers, obtained with artificial intelligence (AI) and digital image analysis (DIA) computer programs, as well as with the projection microscope (PM).

Statistics		Regression type	
		Geometric functional mean	Linear Difference vs average
Regression between AI-Based computer Software and PM			
– Estimated intercept			3.52
– Standard Error of the intercept			0.77
– Significance of the intercept			***
– Estimated slope		0.97	0.03
– Standard error of slope		0.01	0.01
– Significance of the slope	t-value	2.46	2.45
	Significance	*	*
– Significance of the correlation	r-value	0.99	0.27
	t-value	84.63	0.016
	Significance	***	*
Regression between DIA-Based computer Software and PM			
– Estimated intercept			-9.66
– Standard Error of the intercept			1.52
– Significance of the intercept			***
– Estimated slope		0.81	0.21
– Standard error of slope		0.01	0.03
– Significance of the slope	t-value	13.84	8.29
	Significance	***	***
– Significance of the correlation	r-value	0.99	0.80
	t-value	39.43	<0.001
	Significance	***	***

4. Discussion

The procedure to determine the %MedFib that relies on the PM took 18.57 min/sample on average and required intensive operator work (Balasingam, 2005). This time found is lower than reported by Shakyawar et al. (2013). The DIA-based software is clearly faster, requiring an average of about 240 s/sample. Nevertheless, the AI-based software is the fastest one by a large margin, since it identifies an average of 19.44 s/sample. Both methods proposed here are shown to be faster (especially the AI-based software) and, moreover, they minimize the fatigue caused by manual operator work.

These algorithms, therefore, are able to provide practical and precise methodologies that require little time for determining the incidence of medullation. However, the software developed require their counterpart hardware, so that it can scan the samples of fibers prepared in a slide. This would allow obtaining the results in real time, as it happens with the Fiber-EC (Quispe et al., 2017) and the OFDA (Baxter et al., 1992) when they evaluate the diameter of fibers (and its variation) of different species of animals.

Before this research, there was no practical, rapid and precise procedure or methodology for determining the alpaca fiber medullation. Although the OFDA100 has the ability to determine the medullation of sheep and mohair fibers (Lupton and Pfeiffer, 1998), unfortunately, to

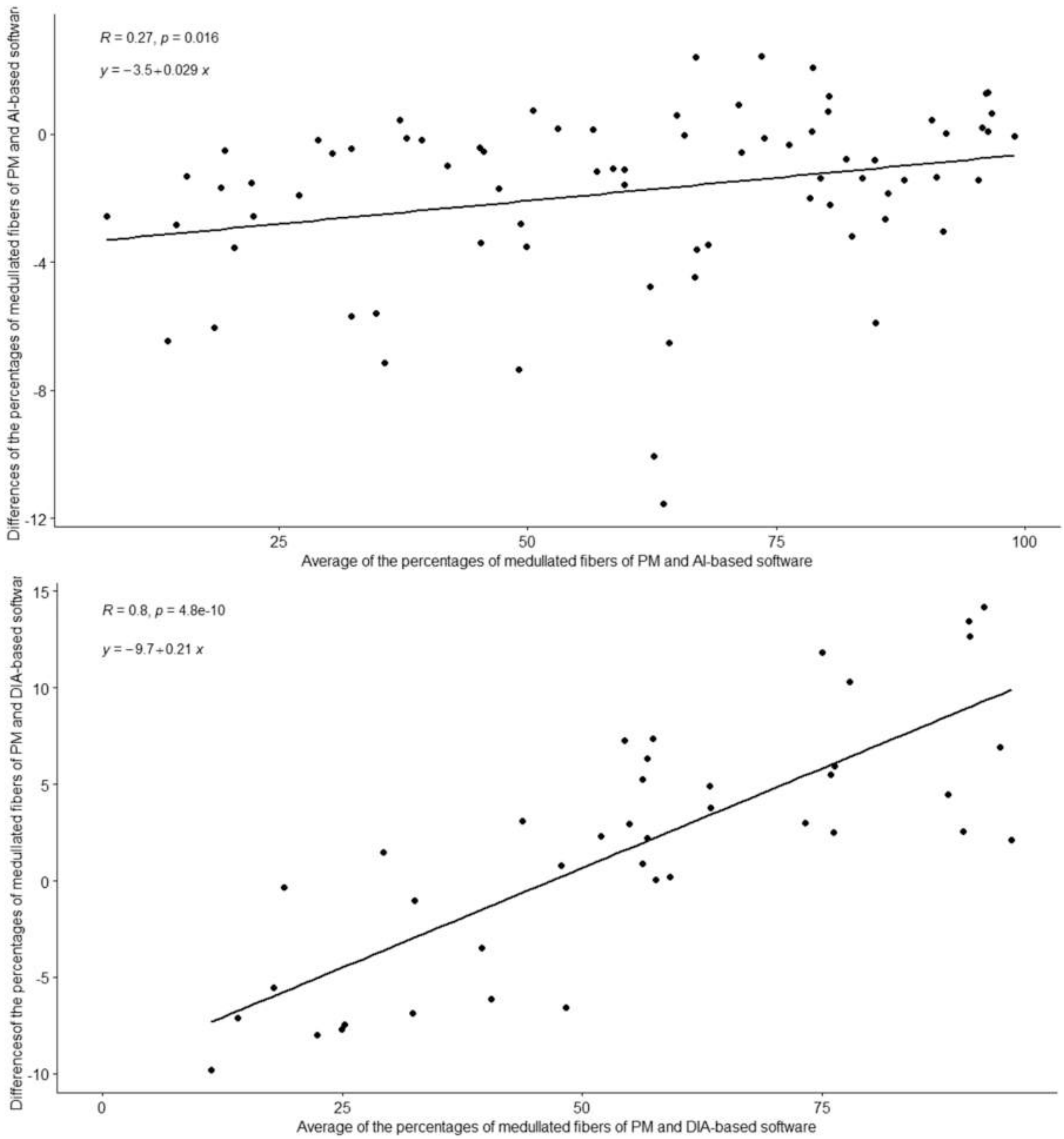

Fig. 2. Scatter plot of the differences against the averages of the percentages of medullated fibers (%MedFib) obtained with the AI- and the DIA-based software with the projection microscope (PM) procedure, on the top and bottom, respectively. Pearson's correlation coefficients and their significance are also shown, as well as the fit-line equation (AI= Artificial Intelligence; DIA: Digital Image Analysis).

date, it does not have accurate results in the evaluation of the medullation of alpaca fibers (Pinares et al., 2018; Torres, 2020). Also, although previously Shakyawar et al. (2013) developed a system based on digital image analysis to measure the diameter and determine the incidence of medullation (considering three categories: heterotopic, kemp and hair), the work was carried out on Chokla native sheep wool and with a minimum number of 42 samples.

On the other hand, currently there is no effort available in the literature on the use of AI for the determination and identification of alpaca fibers by type of medullation. However, this technology has many uses in the recognition of different types of objects, natural language processing, computer vision systems, fraud detection, automatic vehicle handling, disease detection, among others (Krizhenysky and Sutskever, 2012).

The results of the validation indicate that the AI- and DIA-based software show similar %MedFib to those obtained with the PM. The advantage of the two pieces of software presented herein is that they are faster, while being reliable. Additionally, they provide better repeatability, since the results of the PM method cannot be flawlessly replicated even if the same operator evaluates the same sample, varying even more from person to person (Shakyawar et al., 2013).

In addition, the degree of relationship of each of the computer methods with the PM is remarkably strong ($r = 0.99$ and 0.97). This substantiates the ability of the software to recognize fibers with and without medullation. In addition, the results of %MedFib using computer methods are consistent with those obtained by various researchers working with the PM (Pinares et al., 2018; Torres, 2020; Radzick-Rant and Wierckinska, 2021).

The evaluation of bias allowed us to evaluate whether the relationship between the computer methodologies with the PM remains constant at different %MedFib (IWTO, 2017c). The results indicate the existence of a slight bias of AI- and DIA-based software because the regression coefficient (i.e. the slope) is low (0.03 and 0.21, respectively). This low value of b_{xy}, indicates that the relationship between the

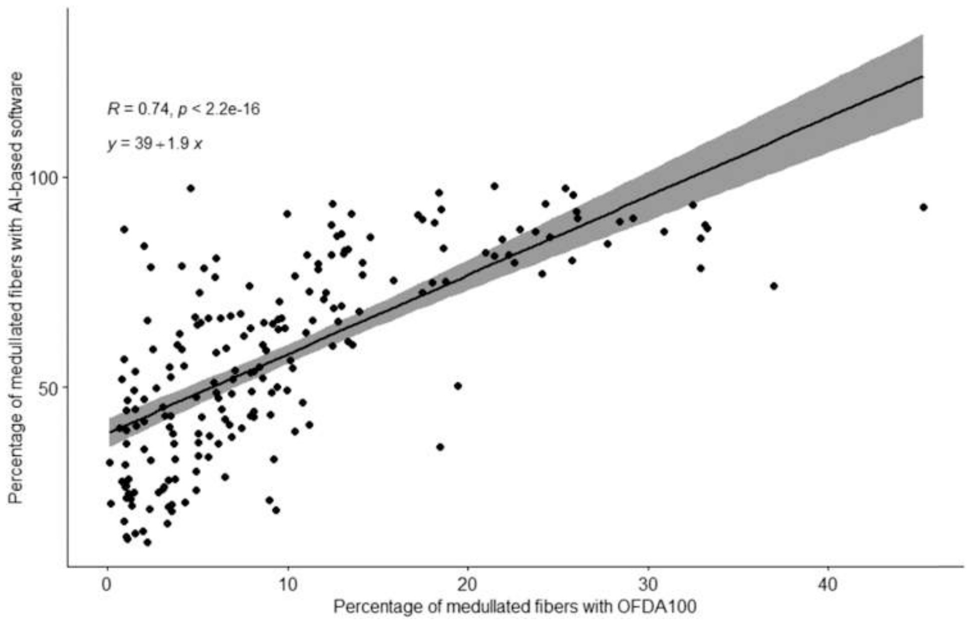

Fig. 3. Scatter graphics of percentages of medullated fibers (%MedFib) assessed with OFDA100 and the AI-based software. It shows Spearman correlation and linear regression equation (AI=Artificial Intelligence).

Table 4
Comparative statistics of the percentage of medullation found by the AI-based software and OFDA 100 in Huacaya (n = 100) and Suri (n = 100) alpaca fibers.

	Mean	Standard error	Minimum	Maximum
Huacaya alpacas				
●AI-based software	54.96^{b}	2.43	13.27	97.89
●OFDA100	10.94^{a}	0.92	0.90	42.25
Suri alpacas				
●AI-based software	39.86^{b}	2.17	2.66	82.35
●OFDA100	8.69^{a}	0.78	0.08	33.17

AI-based software and the PM method would have a maximum variation of 3% when fibers with a high incidence of medullation are evaluated. This could be considered small, given that there is an almost perfect relationship between these methods (99%). Thus, due to the bias effect, the relationship could be reduced down to a 96%. The results of the relationship between the DIA-based software and the PM method are not as good (21%) and, therefore, caution must be observed. The intercept and correlation between the means and differences of the PM compared with the IA-based software is lower than the PM compared with the DIA-based software. The slight bias has no influence on the similarity of the data, nor on the relationship found between them. Likewise, it is reasonable that the least accurate and precise method for the determination of the %MedFib in fibers with high incidence of medullation would be the PM method, since the procedure is quite tedious and strenuous for the operator (Balasingam, 2005; Hunter et al., 2013; Torres, 2020). With the PM method, the operator performs more counts when working with fiber samples of higher %MedFib than when there is less medullation incidence. Therefore, they will spend more time on the computer screen, which could affect their eyes health. Indeed, Computer Vision Syndrome (CVS) has been a recognized as a health condition for over 20 years (Sheppard and Wolffsonhn, 2018).

The results of the %MedFib obtained with the OFDA100 (which are inferior compared to the procedure that uses AI) are concordant with those obtained by Pinares et al. (2018) and Torres (2020). They compare the results obtained with the PM and the OFDA100. However, they did not use alpaca fibers for the validation. Thus, while OFDA100 has been validated in sheep wool (Lupton and Pfeiffer, 1998; Baxter, 2002; IWTO, 2017b), in mohair the results, depending on the sources consulted, vary from accurate to arguable (Lupton et al., 1991; Brims and Peterson, 1994; Turpie and Steenkamp, 1995). Lee et al. (1996) showed that OFDA100 underestimates medullation values in mohair fibers by up to 8%. These considerations suggest that, apparently, the limits of the degrees of opacity used by the OFDA100 for the identification of medullated and non-medullated fibers of alpaca fibers should be revised. Rafat et al. (2007) also indicated that a new definition and calibration of OFDA must be developed to measure medullation in Angora rabbit fibers. Although Angora fibers have a percentage of medullation close to 100% (Berollati et al., 2021), they found a low average incidence of medullation using the OFDA100. In this context, it has been observed that all Angora fibers have at least 1 medulla canal (Chattopadhyaya et al., 2005; Rafat et al., 2007).

The differences of %MedFib observed between AI and OFDA100 could be because OFDA100 was developed to measure incidence of medullated fibers as a function of fiber opacity using wool (specifically, Merino wool), mohair and cashmere (Brims and Peterson, 1994; Balasingam, 2005). The Merino wool has low percentage of medullated fibers (Berollati et al., 2021), and mohair is similar as regards to medullation incidence. Fiber opacity is defined as the ability of a fiber to transmit light perpendicular to the fiber length. It is calculated by summing the light transmitted by the fiber in dark field mode, normalized by the fibre diameter. When measuring medullation with the OFDA100, 80% opacity is generally regarded as representative of normal medullated fiber population (Balasingam, 2005). Fiber opacity is in turn affected by the optical lens effect of the fiber cross section and the effect of a keratin/air interface within the fiber (Brims and Peterson, 1994), and it is determined by the fiber's shape, internal structure, color and surface quality (Baxter, 1998; Hornik, 2012; Maqalika, 2020). Therefore, the OFDA100 do not distinguish between pigmentation versus medullation. In white animal fibers, the main cause of opacity is medullation (hollow fibers). It is thus unlikely to detect medullae of small diameter in thin fibers (Reid et al., 2007), and probably in other types of animal fibers such as alpaca (Pinares et al., 2018; Torres, 2020), Angora rabbit (Rafat et al., 2007), or *Brushtail possum* fibers (Reid et al., 2007), because they have different properties of wool. Alpaca fibers have higher %MedFib, the scales are thinner and the shape and structure are different compared to wool (McGregor and Quispe, 2018).

It is worth noting here than, currently, the OFDA100 has been withdrawn from the market and it has replaced by OFDA2000. However, this new equipment is not adapted to measure the light transmitted by the fiber and, therefore, it cannot measure the incidence of medullated fibers (Hornik, 2012).

5. Conclusion

The software developed (based on both AI and DIA) allows evaluating a large number of alpaca fiber images in a short time, compared to previous efforts. It also allows determining the incidence of medullation. The results show that the proposed software is faster while being reliable and more repeatable. Between the two options proposed, the AI-based software identifies a greater number of fibers in less time. Nonetheless, both pieces of software are faster than the standard traditional procedure (the PM method), which takes a long time and requires intensive operator labor. Although the algorithms proposed would allow the implementation of practical, precise, efficient methodologies that require little time for determining the incidence of medullation, there is still a need to design and develop the counterpart hardware, which will also help to obtain the results in real time.

Conflict of interest

None.

Acknowledgements

We would like to thank the anonymous reviewers for their comments and suggestions that helped to improve the preliminary version of the manuscript. This research did not receive any specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Open Access funding provided by Public University of Navarre, Spain.

References

Arcidiácono, M., Constable, L., Destefanis, E., Vázquez, J., 2014. Determining diameter of animal textile fiber using image processing techniques. In: Proceedings of the XL Latin AmericanComputing Conference (CLEI), Montevideo, IEEE, 1–6.

Aylan-Parker, J., McGregor, B., 2002. Optimising sampling techniques and estimating sampling variance of fleece quality attributes in alpacas. Small Rumin. Res. 44, 53–64. https://doi.org/10.1016/S0921-4488(02)00038-X.

Balasingam, A., 2005. The Definitions of Medullation Threshold Values used by Different Testing Methods to Define an Objectionable Medullated Fibre in Merino Wool. AWTA Limited, New South Wales.

Baxter, B., 2002. Raw-wool metrology: Recent developments and future directions. Wool. Technol. Sheep Breed. 50 (4), 29–38. ⟨https://www.sgs.com/~/media/Global/Documents/Third%20Party%20Documents/Third%20Party%20Technical%20and%20Research%20Papers/TP_M20_General%20Wool%20Metrology%20Issues⟩. accessed 11 March 2021.

Baxter, B., Brims, M., Taylor, T., 1992. Description and performance of the optical fibre diameter analyser (OFDA). J. Text. Inst. 83 (4), 507–526. https://doi.org/10.1080/00405009208631225.

Baxter, B.P., 1998. Some Notes on the Influence of Colour on the Measurement of Medullation by OFDA. International Wool Textile Organisation, Sliver Group, Nice, France.

Berollati, G., Ruiz, L., Cabrera, F.A., Aliaga, J.L., Quispe, M.D., Quispe, E.C., 2021. Evaluación de la medulación de fibras de lanas y fibras especiales de algunas especies de animales. Rev. Inv. Vet. Perú 32 (5), e17639 https://doi.org/10.15381/rivep.v32i5.17639.

Blakeman, N., Lupton, C., Pfeiffer, F., 1988. A sonic digitizer technique for measuring medullation in Mohair. Text.Res. J. 58 (9), 555–556. https://doi.org/10.1177/004051758805800911.

Boguslavsky, A., Botha, A., Hunter, L., 1992. MeasuringMedullation in Mohair with Near Infrared Reflectance Analysis. Text.Res. J. 62 (8), 433–437. https://doi.org/10.1177/004051759206200801.

Botha, A., Hunter, L., 2010. The measurement of wool fibre properties and their effect on worsted processing performance and product quality. Part 1: the objective measurement of wool fibre properties. Text. Prog. 42 (4), 227–339. https://doi.org/10.1080/00405167.2010.486932.

Brims, M., Peterson, A., 1994. Measuring Fibre Opacity and Medullation using OFDA: Theory and experimental results on mohair. IWTO Technical Committee Report. IWTO, New Delhi.

Chattopadhyaya, S.K., Bhaskar, P., Ahmed, M., 2005. Properties of indigenous angora rabbit hair and cotton blended yarns spun using short-staple cotton spining system. Indian J. Fibre Text. Res. 30, 215–217. ⟨http://nopr.niscair.res.in/bitstream/123456789/24680/1/IJFTR%2030%282%29%20215-217.pdf⟩. accessed 07 march 2022.

Cottle, D., Baxter, B., 2015. Wool metrology research and development to date. Text. Prog. 47 (3), 163–315. https://doi.org/10.1080/00405167.2015.1108543.

Cruz, A., Morante, R., Gutiérrez, J., Torres, R., Burgos, A., Cervantes, I., 2018. Genetic parameters for medullated fiber and its relationship with other productive traits in alpacas. Animal 13 (7), 1358–1364. https://doi.org/10.1017/S1751731118003282.

Czaplicki, Z., 2012. Properties and structure of Polish Alpaca Wool. Fibres Text. East. Eur. 1 (90), 8–12.

Deng, Z., Ke, W., 2010. A new measuring method of wool fiber diameter based on image processing. IEEE, Dalian, China, pp. 587–590.

Frank, E.N., Castillo, M., Prito, M., Adot, O., 2014. Fiber-based components determining handle and skin comfort in fabrics made from dehaired and non dehaired llama fiber. Int. J. Appl. Sci. Technol. 4, 51–66. ⟨http://www.ijastnet.com/journals/Vol_4_No_3_May_2014/7.pdf⟩. accessed 20 November 2021.

Gupta, N., Arora, R., Verma, G., 1981. An assessment of the characteristics of medullated and non-medullated wool fibres. Indian J. Text. Res. 6, 92–95.

Hack, W.M., Ponzoni, R., Judson, G., Carmicheal, I., Hubbard, D., 1999. Australian Alpaca Fibre: Improving, Productivity and Marketing. Rural Industries Research and Development Corporation Research Paper Series, Melbourne.

Hornik Fibertech, 2012. OFDA100. A World Breakthrough inf Fibre Mesurement Technology, OFDA100, chefistrasse, Switzerland. ⟨https://www.hornik.cc/ofda_100.php⟩. (Accessed 7 March 2022).

Hunter, L., Smuts, S., Botha, A., 2013. Characterizing visually objectionable and nonobjectionable medullated fibers in Mohair. J. Nat. Fibers 10, 112–135. https://doi.org/10.1080/15440478.2013.763483.

IWTO, 2017a. IWTO-8-2011 Method of Determining Fibre Diameter Distribution Parameters and Percentage of Medullated Fibres in Wool and Other Animal Fibres by the Projection Microscope, 19. International Wool Textile Organization, Brussels.

IWTO, 2017b. IWTO-57-200 Determination of Medullated Fibre Content of Wool and Mohair Samples by Opacity Measurements Using an OFDA. International Wool Textile Organization, Brussels.

IWTO, 2017c. IWTO-0-2012 Appendix B Presentation of Supporting Technical Data. Introduction to IWTO Specifications. Procedures for the Development, Review, Progression or Relgation of IWTO Test Methods and Draft Test Methods. International Wool Textil Organization, Brussels.

Krizhenysky, A., Sutskever, I.H., 2012. ImagenNet Classification with deep convolutional neural networks. In: Proceedings of the Twenty Fifth International Conference on Neural Information Processing Systems, Toronto, NIPS'12, (pages 1: 1097-1105).

Lappage, J., Bedford, J., 1983. WRONZ, Christchurch: Report N°107.

Lee, J., Maher, A., Frampton, C., Ranford, S. 1996. Comparison of medullation in the same fiber sites using OFDA. IWTO Technical and Standards Committee, Special Topics Group, (Rep. No. 14), Capetown, South Africa.

Lupton, C., Pfeiffer, F., 1998. Measurement of medullation in wool and mohair using an optical fibre diameter analyser. J. Anim. Sci. 76, 1261–1266. https://doi.org/10.2527/1998.7651261x.

Lupton, C., Pfeiffer, F., Blakeman, N., 1991. Medullation in mohair. Small Rumin. Res. 5, 357–365. ⟨https://agrilifecdn.tamu.edu/sanangelo/files/2011/11/R-116-Medullation-in-mohair.pdf⟩. accessed 16 december 2021.

Maqalika, P.E., 2020. Database and guide for Lesotho wool and mohair production andquality. Thesis for the Philosphy Doctor of Textile Science. Nelson Mandela University, South Africa, 255.

Martinez, Z., Iñiguez, L., Rodríguez, T., 1997. Influence of effects on quality traits and relationships between traits of the llama fleece. Small Rumin. Res. 24, 203–212. https://doi.org/10.1016/S0921-4488(96)00925-X.

McGregor, B., 1997. The quality of fiber grown by Australian Alpacas: part I. The Commercial Quality Attributes and Value of Alpaca Fiber. Australia Alpaca Association, Melbourne.

McGregor, B., Ramos, H., Quispe, E., 2011. Variation of fibre characteristics among samples sites for Huacaya alpaca fleeces from the High Andes. Small Rumin. Res. 102 (2–3), 191–196. https://doi.org/10.1016/j.smallrumres.2011.07.016.

McGregor, B., Quispe, E., 2018. Cuticle and cortical cell morpjology of alpaca and rare animal fibres. J. Text. Inst. 109 (6), 767–774. https://doi.org/10.1080/00405000.2017.1368112.

Pinares, R., Gutiérrez, G., Cruz, A., Burgos, A., Gutiérrez, J., 2019. Variabilidad fenotípica del porcentaje de fibras meduladas en el vellón de alpaca Huacaya. Rev. Investig. Vet. del Peru 30 (2), 699–708. https://doi.org/10.15381/rivep.v30i2.16098.

Pinares, R., Gutiérrez, G., Cruz, A., Morante, R., Cervantes, I., Burgos, A., Gutiérrez, J., 2018. Heritability of individual fiber medullation in Peruvian alpacas. Small Rumin. Res. 165, 93–100. https://doi.org/10.1016/j.smallrumres.2018.04.007.

Qi, K., Lupton, C., Pfeiffer, F., Minikhiem, D., Kumar, N., Whittaker, A., 1995. Automatic image analysis system for objective measurement of animal fibers. Sheep Goat Res. J. 11 (2), 71–77. ⟨http://sanangelo.tamu.edu/files/2011/11/R128-Automatic-Image-Analysis-System-for-Objective-Measurement-of-Animal-Fibers.pdf⟩. accessed 15 August 2021.

Quispe, E., Chipa, L., Pinares, E., 2015. Analysis of the productivity and economics of manual dehairing of Chaku llama (*Lama glama*). Arch. De. Zootec. 64 (246), 191–198. https://doi.org/10.21071/az.v64i246.397.

Quispe, M., Benavidez, G., Sauri, R., Bengoechea, J., Quispe, E., 2017. Development and preliminary validation of an automatic digital analysis system for animal fibre analysis. South Afr. J. Anim. Sci. 47 (6), 822–833. https://doi.org/10.4314/sajas.v47i6.10.

R Core Team. 2016, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-9000-51-07-0. ⟨https://www.r-project.org/⟩.

Rafat, S.A., de Rochambeau, H., Brims, M., Thébault, R.G., Deretz, S., Bonnet, M., Allain, D., 2007. Characteristics of Angora rabbit fiber using optical fiber diameter analyzer. J. Anim. Sci. 85, 3116–3122.
Radzick-Rant, A., Wierckinska, K., 2021. Analysis of the wool thickness and medullation characteristics based on sex and color in a herd of alpacas in Poland. Arch. Anim. Breed. 6, 157–165. https://doi.org/10.5194/aab-64-157-2021.
Reid, T.C., Causer, S.M., Urquhart Woods, R.A., 2007. Variation of fibre characteristics important in processing, over the body of Australian brushtail possum (Trichosurus vulpecula). Proc. N. Z. Soc. Anim. Prod. 67. In: ⟨http://www.nzsap.org/system/files/proceedings/2007/ab07064.pdf⟩. accessed 01 march 2022.
Shakyawar, D., Kadam, V., Surya, A., Ahmed, A., Pareek, P., Temani, P., 2013. Precise measurement of wool fibre diameter using computerized projection microscope. Indian J. Small Rumin. 19, 190–192. ⟨https://krishi.icar.gov.in/jspui/bitstream/123456789/47064/1/13-Precise%20Measurement%20of%20Wool%20Fibre%20Diameter%20Using%20Computerized%20Projection%20Microscope.pdf⟩ (Accessed 10 October 2021).
Sheppard, A.L., Wolffsonhn, J.S., 2018. Digital eye strain: prevalence, measurement and amelioration. BMJ Open Ophthalmol. 3, e000146 https://doi.org/10.1136/bmjophth-2018-000146.
Shelton, M., 1995. Angora goat and mohair production. Anchor Publishing Company, San Angelo, USA, p. 113 (pages).
Tarqui, N.I., 2008. Evaluación de la calidad yrendimiento de fibra clasificada y descerdada de vicuña (Vicugna vicugna)criadas en condiciones de semicautiverio en Patacamaya (Degree thesis), Universidad Mayor de San Andrés, La Paz, Bolivia, 81.
Torres, R., 2020. Tasa de medulación de fibra dealpaca (Vicugna pacos) mediante la comparación del medulómetro y el OFDA 100. Thesis of M.Sc. Universidad Católica de Santa María, Arequipa.
Turpie, D., Steenkamp, C., 1995. Objective measurement of "objectionable" medullated fibres in commercial mohair tops using an optical fibre diameter analyser (OFDA): An Introductory Study. IWTO, Harrogate.
Ultralytics, 2020. YOLOv5. ⟨https://github.com/ultralytics/yolov5⟩. (Accessed 10 June 2021).
Villarroel, J., 1963. Un estudio de la fibra de alpaca. An. Cient. 1 (3), 246–274.
Wang, H., Liu, X., Wang, X., 2005. Internal structure and pigment granules in coloured Alpaca fibers. Fibers Polym. 6, 263–268.
Wang, X., Wang, L., Liu, X., 2003. The Quality and Processing Performance of Alpaca Fibres. RIRDC Publication, Melbourne, Australia accessed 15 june 2021. ⟨https://www.agrifutures.com.au/wp-content/uploads/publications/03-128.pdf⟩.
Wood, E., 2003. Textile Properties of Wool and Other Fibres. Wool. Technol. Sheep Breed. 51 (3), 272–290.