

A RAPID AND NON-DESTRUCTIVE IDENTIFICATION OF ANIMAL HAIRS USING ATR-FTIR AND CHEMOMETRICS: A PROOF-OF-CONCEPT FOR WILDLIFE FORENSIC APPLICATIONS

Saptarshi RAO¹ , Deepali JAIN¹, JR GAUR¹, Rajesh VERMA²

¹ School of Behavioral Sciences and Forensic Investigations, Rashtriya Raksha University Gandhinagar, Lavad-Dahegam, Gujarat, India

² Directorate of Forensic Services, Himachal Pradesh, Shimla Hills, Junga, Himachal Pradesh, India

Abstract

Wildlife crime is a significant threat to biodiversity and can have serious ecological, economic, and social impact. Skin, horns, claws, antlers, and virtually all parts of an animal's body are utilized in illegal trade. Animal hair is invariably found as physical evidence in wildlife crimes pertaining to mammals. It is also found in wildlife crimes in the form of illegal artifacts, or as circumstantial evidence suggesting the involvement of crime against animals. DNA typing methods are widely applied for species identification but are sometimes unreliable when the sample is highly degraded or mixed with other items. Hair is commonly analysed by microscopic techniques; however, it lacks statistical confidence in identification when the sample size is small and the results are somewhat subjective in nature. Here, we investigate the role of attenuated total reflection Fourier transform-infrared (ATR-FTIR) spectroscopy in analysing the spectra obtained from the hair of two distant species of Indian blackbuck (*Antelope cervicapra*) and Hanuman langur (*Semnopithecus entellus*) in combination with a suitable chemometric model, i.e., PCA (principal component analysis) and PLS-DA (partial least squares discriminant analysis). This is an alternate non-destructive method for the distinction of the multiple spectra. PCA plot showed the grouping to some extent; however, PLS-DA analysis resulted in the correct segregation of both species. Additionally, this model was validated by 6 unknown hair samples of both species, resulting in a 100% accuracy. The model's sensitivity and specificity were also tested and calculated to be 1. Hence, the potential of ATR-FTIR spectroscopy is demonstrated by its speed, non-destructive examination, and minimal or no sample preparation. It can complement the present microscopic and DNA-based techniques.

Keywords

Forensic science; Wildlife forensics; ATR-FTIR spectroscopy; Species differentiation; Chemometrics.

Received 11 March 2024; accepted 14 May 2024

1. Introduction

Hair fibers are present everywhere and found invariably in wildlife crimes pertaining to mammals. After obtaining an animal hair sample, the main aim of investigators is to determine the species [1]. In wildlife forensics investigation and biology, analysis of animal hair could be a useful technique for locating unlawful

trade, slaughter, and poaching of animals, especially those that are endangered. Hair specimens are highly resistant to decomposition and hence permit species identification even from aged samples [2]. Different species have a variety of features such as color, texture, scale patterns, medullary index and their patterns, pigmentation, etc. that can be examined microscopically for species identification. The chemical analysis

of hair can indicate the diet, environmental condition, stress, and geographic origins. This information can be crucial in understanding the movement patterns, habitat preferences, and illegal trade routes associated with wildlife crimes [3].

According to the UNEP-INTERPOL report: The Rise of Environmental Crime, the annual value of the illegal wildlife trade might hit USD 20 billion (Wildlife Crime, 2022). With an annual trade of 7–23 billion USD and as per the United Nations Office on Drugs and Crime (UNODC) it is considered the fourth most illegal trade after arms, drugs, and human trafficking. Less than 30,000 rhinos of all species (including over 18,000 white rhinoceros from Africa) are still alive in the globe, whereas Indian forests are home to close to 3600 rhinos. Rhinoceros populations have dramatically declined over the past 25 years, and their numbers remain still low. Over a century, the number of elephants worldwide has dropped dramatically from 10,000,000 to just 500,000, of whom 30,000 live in India (with 50% of them in the jungles of Karnataka, Assam, and Kerala) [4]. In yet another report by TRAFFIC, over 1000 pangolins were poached between 2018 and 2022 in India. Live pangolins were found in 50% of seizures, whereas pangolin scales were found in 40% [5]. Gorillas are frequently trapped and killed in traps designed for other animals or stolen and sold as bushmeat. Meat is sold for 5 to 6 US dollars per piece. Due to conflict caused by people, habitat loss, and poaching, scientists claim that there are only approximately 600 gorillas alive. Due to overharvesting and illegal trade, tens of thousands of turtles die every year. In several communities, sea turtles are hunted for their meat, eggs, leather, and shells [6].

Animal hair has been studied at the microscopic level extensively [2, 7, 8, 9, 10, 11, 12]. However, there is a concern that mere microscopic analysis of hair might be subjective and does not have any statistical foundation attached to it [13]. Two experts may have somewhat different opinions or might describe the characteristics using separate terminologies [14]. Besides, the hair of both species is different in colour. But in real case scenarios, they can be found in a highly degraded state that could hinder their identification accurately. Based on Locard's principle of exchange, whenever two objects come in contact with each other, an exchange takes place. If hair is found in cases of illegal trapping or during a struggle, it can serve as potential evidence.

Currently, the most sought-after method for identifying species is DNA-based identification. Although DNA is unique to each species and can be used to distinguish between species, it is important to be

aware that many hair strands lack sufficient root tissue, which could make DNA analysis difficult [15]. In this purview, a method that offers a rapid, reliable, and non-destructive examination of hair samples, complements the microscopic analysis with a certain level of statistical assurance will greatly prove beneficial in forensic investigations.

Attenuated total reflection Fourier transform-infrared (ATR-FTIR) spectroscopy has many appealing qualities and is growing in popularity in the analytical and biomedical domains due to its non-destructive, dependable, and rapid nature [15]. The technique is widely used in material science, chemistry, biology, and forensic science. In wildlife forensics, ATR-FTIR spectroscopy combined with chemometrics is an effective tool for identifying illegally traded wildlife products, such as horns, shells, and corals, based on their chemical composition. A select few studies demonstrate how to distinguish between infrared spectra from different types of hair by using chemometric methods. Espinoza et al. investigated the use of horizontal ATR-FTIR for the study and identification of elephant and giraffe tail hair, which is frequently utilized to make native jewellery (such as bracelets, finger rings, and earrings) in the wildlife trade [16]. Another research group visually discriminated synthetic fibre from human hair and then distinguished human hair from animal hair using chemometric models of ATR-FTIR spectroscopic data [1]. PLS-DA model presented a 100% predictability of whether a hair specimen was human or non-human. Apart from these, Xu and team analysed six different animal fur types and compared 4 kinds of chemometric tools, SIMCA, PCA-LDA, PLS-DA, LS-SVM. Out of them, PLS-DA and LS-SVM proved to be 100% effective techniques [17]. Gao used FTIR spectroscopy to examine the straight guard hair of the Indian leopard cat (*Prionailurus bengalensis*) and golden cat (*Catopuma temminckii*) [18]. An evident peak was seen at 648 and 654 cm^{-1} in golden cat but not in leopard cat. The second derivative spectra of the golden cat had strong peak intensity while that of the leopard cat had weak peak intensity. Of late, Sharma et al. [15] presented a novel technique to differentiate *shahtoosh* wool from *pashmina* wool. Recently, Guleria et al. [19] successfully differentiated between the feather samples of 12 species of peacock (*Galliformes*) using PCA-LDA with 100% accuracy.

It is a known fact that conventional methods have been used for species identification but their destructive nature is still a concern [15]. Therefore, ATR-FTIR spectroscopy is very adaptable from the perspective of forensics and can be an addition to microscopic analysis, DNA analysis, and other techniques that

are already used. As far as the author's knowledge, no significant work has been reported on the analysis of hair from different species of wild animals using ATR-FTIR. Previous studies implemented domestic species which are not of much importance in forensic wildlife investigations. Therefore, this preliminary study aimed to differentiate between the hair spectra of 2 distantly related species of Indian blackbuck (*Antilope cervicapra*) and Hanuman langur (*Semnopithecus entellus*) via ATR-FTIR spectroscopy with a suitable chemometric technique, i.e. PCA and PLS-DA. This study aimed to demonstrate that although the hair looks similar morphologically and spectra wise, they can be differentiated using chemometric models. This study can serve as a first-hand guide to utilizing ATR-FTIR for examining illegal wildlife artifacts like hair, blood, bones, or teeth.

The main goal of using chemometric tools or multivariate data analysis with ATR-FTIR spectroscopic methods is to reduce large datasets so that results can be easily interpreted.

2. Materials and methods

2.1. Hair sample collection

Reference hair samples from the dorsal part of Indian blackbuck and Hanuman langur were collected from Indroda Nature Park, Gandhinagar (23°14'15.2160"N and 72°38'52.0116"E). Ethical clearance was obtained from the Principal Chief Conservator of Forests and Chief Wildlife Warden, Gujarat State (WLP/RTC/28/C/57-59/2023-24). Details are enumerated in the table below.

Table 1
Detailed description of samples

Common name	Scientific name	Number of individuals	Number of hair strands per species
Indian blackbuck	<i>Antilope cervicapra</i>	4	25
Hanuman langur/ Gray langur	<i>Semnopithecus entellus</i>	4	25

2.2. Sample preparation and spectral collection

Each hair sample was sonicated in water for 10 minutes, then washed with isopropyl alcohol for 10 minutes to reduce the possibility of contamination

[16]. A single hair strand was placed right on the instrument's crystal cell. Each hair strand was gently pressed with the ATR pressure anvil to at the midpoint to ensure sufficient contact between the hair sample and the crystal surface before the spectrum was captured. Before analyzing any new hair samples, the crystal of the ATR was thoroughly wiped with ethanol and a background spectrum of the empty ATR cell was gathered under the same conditions. By examining three duplicates of each collected hair strand, reproducibility was examined. To examine the homogeneous distribution of components within samples, a homogeneity test was carried out on all of the hair strands by examining the same strand from three separate places (proximal, medial, and distal). However, valid data could also be acquired from a single hair when its diameter filled at least 25% of the sampling window (hair diameter 0.5 mm), even though the best spectra were obtained when the hairs entirely covered the sampling window (2 mm) [16].

2.3. Instrumentation

The samples were examined using a Fourier transform-infrared spectrometer and an attenuated total reflectance (ATR) platinum diamond accessory (Bruker alpha II ATR; Billerica, MA) containing ZnSe crystal plate within the mid-infrared range (MIR), i.e. 4000–600 cm^{-1} at a resolution of 4 cm^{-1} . A total of 24 scans were taken for each spectrum. The spectral data was collected using Opus software (version 8.1). Deuterated lanthanum α alanine doped triglycine sulphate detector (RT-DLaTGS) was utilized.

2.4. Pre-processing of raw data matrix

The primary goal of the data pre-treatment is to reduce any noise or interference in each spectrum. This enables us to focus only on the variations due to the chemical constituents of the samples and not from external sources. In the present work, baseline correction was applied to the raw spectra, and normalization of the data was done using SNV (standard normal variate) which removes slope variation from the spectrum [20]. SNV transformation works by decreasing the mean and scaling the deviation.

2.5. Applied chemometric models

A subfield of analytical chemistry called chemometrics combines statistical and mathematical techniques to infer information from chemical data. The PCA (principal component analysis) and PLS-DA

(partial least squares discriminant analysis) model was created to distinguish total 50 spectra (25 hairs per species) obtained from Indian blackbuck and Hanuman langur in the mid-infrared range, i.e. 4000–600 cm^{-1} . SNV transformation was performed using the Unscrambler X 10.4 (CAMO Software AS, Oslo, Norway) software. R-software which runs through RStudio IDE was used for PCA and PLS-DA. The package mixOmics was used that has been specifically developed for classifying large data sets.

2.5.1. Principal component analysis (PCA)

By breaking down complex data sets into a new collection of variables known as principal components, which can accurately capture the whole range of variance in the data, PCA decreases the dimensionality of the data. Unsupervised dimensionality-reduction or bilinear modelling tools like PCA are well-known for helping to show important data in more understandable ways. Principal components are the variables that are utilized in PCA to determine the relationship among highly correlated variables. It uses the eigenspace method, sometimes referred to as the Karhunen-Loeve transform, and is reflected as a crucial tool to illustrate the important correlations in the chosen samples. Score plots are the most popular visual representation of PCA results. It illustrates the key distinctions and parallels between the datasets, making it simple to understand how the findings should be interpreted [21].

2.5.2. Partial least squares discriminant analysis (PLS-DA)

PLS-DA is a well-known linear classification technique that combines PLSR's (partial least square

regression) distinguishing characteristics with the classification method's ability to discriminate. The flexible PLS-DA approach, which finds the latent variables with the highest covariance with the Y-variables and compares them with the dataset's dummy values, is built on the foundation of the PLSR methodology for calibration. If this dummy value juxtaposes with the result, the sample falls into that category or, in another manner, belongs to a separate classification group [22, 26, 27, 28]. It generates numerous results since it may be used for discriminative variable selection, descriptive modelling, and predictive modelling all at once.

2.6. Validation

For method validation, six unidentified samples from one individual of each species were examined to determine how well the model worked. The analyst was unaware of these samples, and they were not included in the training data set. It was then observed whether they classified themselves in the right category or not.

3. Results and discussion

3.1. Spectral analysis of training dataset

Figure 1 represents the overlaid hair spectra of both species. Chemically, hairs are made up of 65–95% proteins; as a result, amide peaks linked to keratins dominate the peaks of hairs [15].

The different peaks and their corresponding frequencies are enumerated in Table 2.

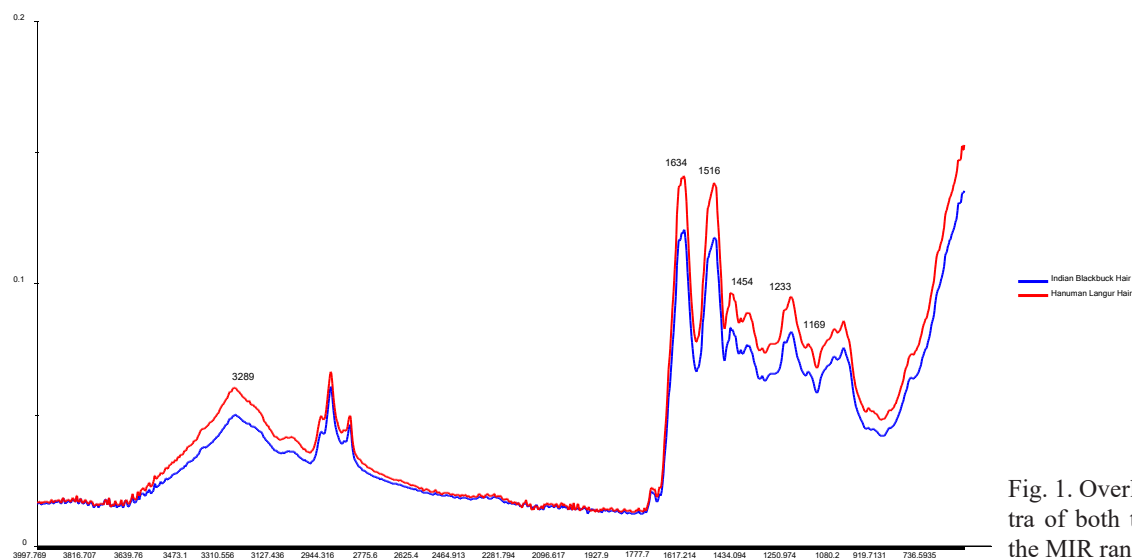


Fig. 1. Overlaid hair spectra of both the species in the MIR range.

As seen in Figure 1, the fingerprint region of the spectra is similar. There is one significant difference seen in the region 3400–2800 cm^{-1} . Only one prominent peak at 3289 cm^{-1} in the monkey hair may be due to H-bonded O-H stretching and N-H symmetric stretching (water and hydroxyl group). Whereas there are two peaks seen in the blackbuck hair spectra at 2921 and 2851 cm^{-1} respectively. These peaks correspond to C-H stretching having medium intensity. Further, two prominent peaks in the hair spectra of both species exist. The 1634 and 1516 cm^{-1} peaks correspond to amide I and II, respectively. The rest of the peaks were characterized as weak due to their low intensity. The peak at approximately 1454 cm^{-1} is due to both the anti-symmetric deformation of CH_3 groups and the scissoring mode of CH_2 groups, while the peak at 1394 cm^{-1} is due to amino acid groups or deformation of the methyl groups.

The spectra were analyzed visually, followed by PCA and PLS-DA methods. The homogeneity test was carried out on all the obtained hair strands by analyzing the same strand from three distinct positions (left, middle, and right), and reproducibility was assessed by examining three replicates of each collected hair strand. Contrary to what was anticipated, no discernible change was found upon direct visible inspection based on the overlay spectra as shown in Figure 1. Both spectra are almost similar and exhibit nearly identical values in wavenumbers with a slight difference, even though the species are distantly related. It is quite difficult to differentiate the spectra solely on visual inspection. It is subjective in nature and hence chemometric models were required to differentiate between them and obtain necessary information.

3.2. Principal component analysis

The pre-processed data was fed into the R-software for PCA analysis. Both the categories of hair samples were balanced in number to ensure the reliability of the results, i.e. 25 hair samples of blackbuck and monkey each. Figure 2 represents a 2D plot of the hair samples which represent the segregation of samples to some extent. Here PC1 explained 62% of the variation and PC2 accounted for 25% variation in the dataset. In total, a cumulative variation of 87% was achieved using the first two PCs. To track the number of PCs that could provide the greatest variance, a scree plot was also created. As can be seen from Figure 2, PCA gives a qualitative idea about the number of classes but does not classify the samples into different classes.

Since PCA is an unsupervised technique and is used for exploratory purposes, PLS-DA analysis was performed for better discrimination and prediction of the unknown samples.

3.3. Discrimination using the PLS-DA model

When compared to the visual comparison and PCA, the primary goal of employing the PLS-DA technique was to establish reliable and satisfactory categorization between the two types of hair samples. The PLS-DA model was built using the training set of 50 samples. To distinguish and categorize the Indian blackbuck hair samples and Hanuman langur hair samples, the model was created with two required latent variables. The PLS-DA approach was applied to the normalized data in R by using the mixOmics package. Large data sets can be handled efficiently by this software, which also allows for customizable sample utilization [23]. Figure 3 demonstrates a 2D plot of the class separation using PLS-DA. Both hair types were

Table 2
Assignment of prominent peaks of animal hair [15] shown in Figure 1

Peak (cm^{-1})	Band assignments	Peak intensity	Functional group
3289	Symmetric N-H stretching in 1° amide	Medium	Amide A
2921, 2851	C-H stretching	Medium	Amide A
1634	C=O stretching, H-O-H bending	Strong & most intense peak	Amide I
1516	N-H deformations in 2° amide	Strong & most intense peak	Amide II
1454	Asymmetric CH_3 bending	Weak	Amide II
1394	Symmetric CH_3 bending	Weak	Amino acid
1233	N-H bending	Weak	Amide III
1169	C-O symmetric stretching	Weak	Cystine oxides

appropriately categorized into their respective groups, as can be seen from both figures, without any misclassifications. When 10 components were selected, the balanced error rate was computed to be 0, as illustrated in Figure 4 (left). The receiver operating characteristic (ROC) curve for two components is shown in Figure 4 (right), along with the outstanding classification model's AUC value of 1.

3.4. Validation

Figure 5 displays the grouping of the unidentified samples, which are represented by black dots, into the appropriate classes. The unknown samples with their predicted class are given in Table 3.

Table 3
Unknown samples prediction

Sample ID	Actual class	Predicted class	Prediction
U1	Blackbuck	Blackbuck	correct
U2	Blackbuck	Blackbuck	correct
U3	Blackbuck	Blackbuck	correct
U4	Langur	Langur	correct
U5	Langur	Langur	correct
U6	Langur	Langur	correct

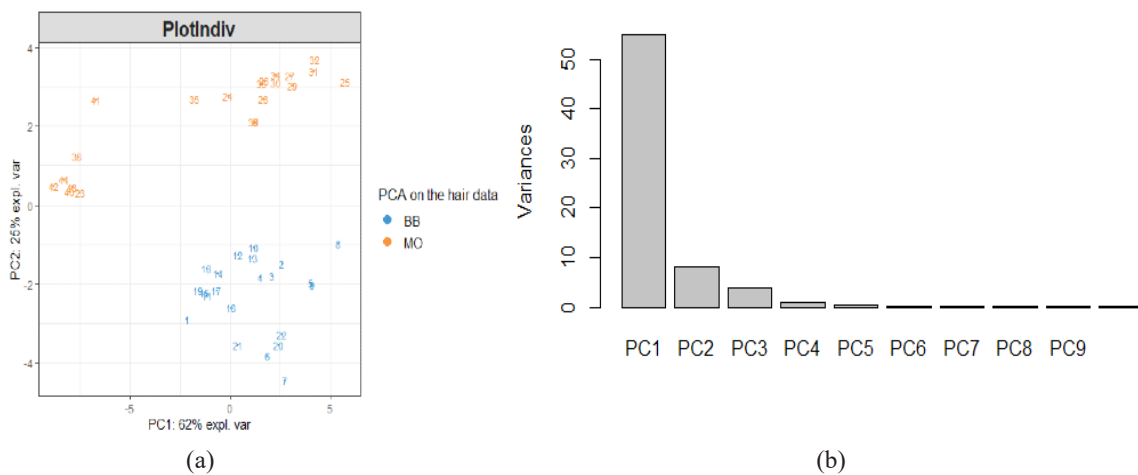


Fig. 2. (a) PCA score plot of Indian blackbuck (BB) and Hanuman langur (MO) hair samples; (b) scree plot of the percentage of explained variance vs. the principal components.

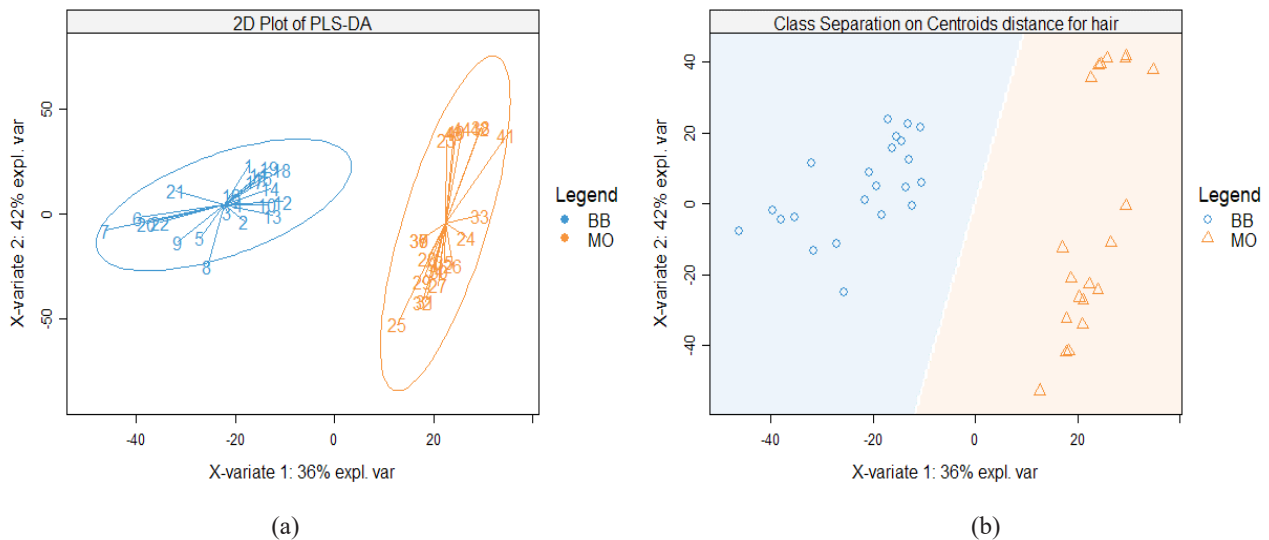


Fig. 3. (a) 2D PLS-DA plot to differentiate Indian blackbuck (BB) and Hanuman langur hair (MO); (b) prediction areas in PLS-DA.

A confusion matrix was also made, as shown in Table 4, to test the suggested model’s sensitivity and specificity. The number of true positive predictions is divided by the total number of true positive and false negative predictions to determine sensitivity [24]. The sensitivity, in this case, was determined to be 1.0. By dividing the total of false positives and true negatives by the number of true negatives, specificity is determined. Specificity was also found to be 1.0. The accuracy determined by the confusion matrix was also found to be 1.0, or 100%.

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \times 100$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \times 100$$

$$\text{Accuracy} = \text{TP} + \text{TN}/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \times 100$$

Table 4
Confusion matrix table

Experimental class	Calculated class	
	Blackbuck	Langur
Blackbuck	7	0
Langur	0	7

Discrimination and identification of species are very important for a variety of forensic applications. In this research, ATR-FTIR spectroscopy and chemometrics are utilized for the discrimination of the hair of wildlife species. An attempt has been made to identify and differentiate these two distantly related species based on the proposed technique. Since the visual analysis was prone to errors and was a time-consuming task, chemometric tools needed to be applied for

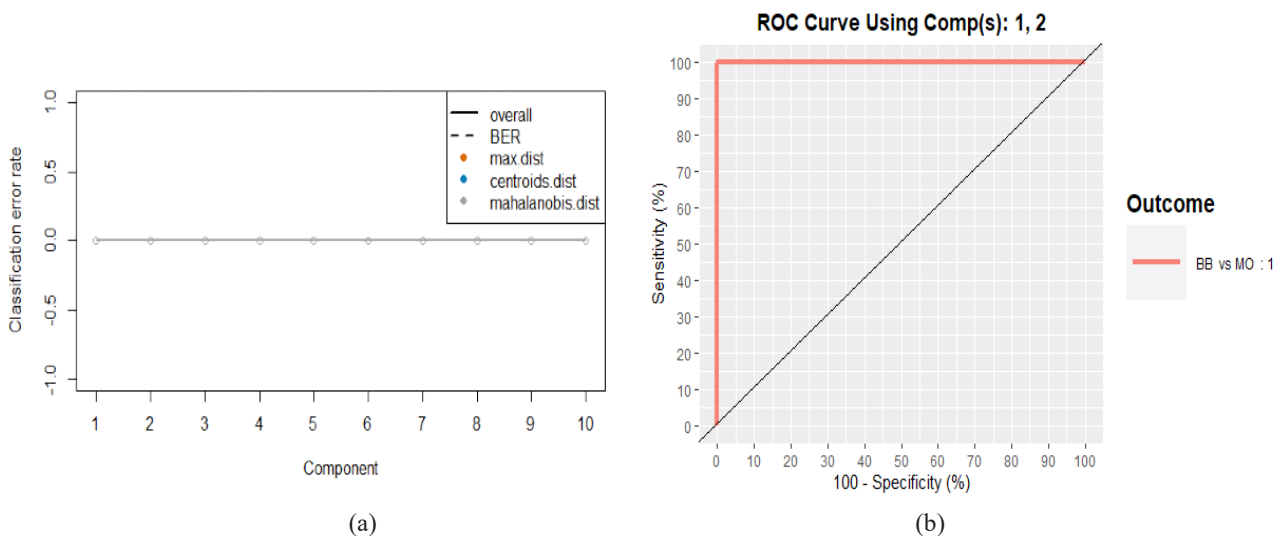


Fig. 4. (a) performance of PLS-DA model on 10 components; (b) ROC curve for PLS-DA.

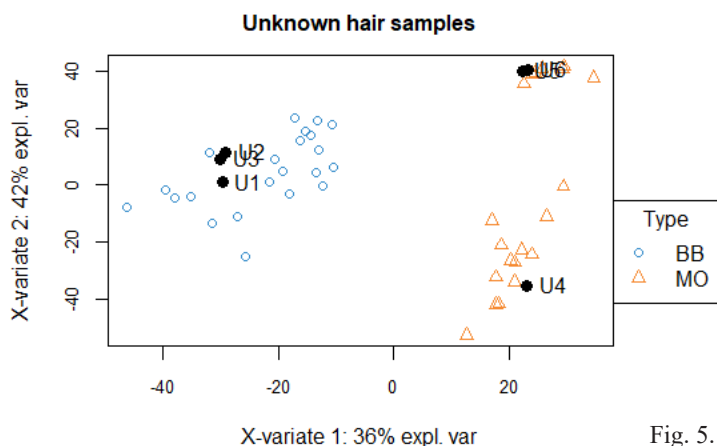


Fig. 5. Prediction of unknown samples by PLS-DA.

better examination. PCA showed a cumulative variation of 87%. It gave discrimination to some extent as well as provided qualitative information about the samples. PLS-DA correctly differentiated both categories of samples with an accuracy rate of 100%. The sensitivity and specificity of the model were found to be 1.0 which signifies that PLS-DA provides for an excellent distinguishing model. For blind testing, 6 unknown samples were tested using the constructed PLS-DA model. All 6 samples were correctly categorized into their respective groups with 0 misclassifications. A summary of all the related studies in this area has been given in Table 5.

4. Conclusion

With a small sample size (just two species), this study gives a preliminary method for hair differentiation and offers a prospective tool for this kind of discrimination. The authors also conclude that this method is not a substitute for DNA-based analysis, but complementary to it. Due to ethical considerations, the sample size is less in the current work. However, larger, more thorough investigations with a larger sample size are needed to fully investigate and test the possibilities of ATR-FTIR spectroscopy. Besides, if hair evidence is found in a highly degraded form

in counterfeit items, then this method can be explored for identification whenever DNA typing may not yield satisfactory results. If the hair of two distant species can be differentiated, then hair artifacts of closely related morphologically similar species can also be differentiated and identified using this approach. The authors also performed this study to investigate this methodology further using hair samples of different species. At present, commercially available portable ATR-FTIR spectrometers can also be utilized for on-field investigations. Apart from hair, experiments can also be undertaken on specimens like animal teeth and bones that are frequently encountered in illegal trade using this methodology.

Thus, the present study concluded that ATR-FTIR spectroscopy and chemometrics present a robust, non-destructive method for species identification in wildlife forensic investigations.

Ethical approval

All procedures involving animals were in compliance with the European Community Council Directive of 24 November 1986, and ethical approval was granted by the Principal Chief Conservator of Forests & Chief Wildlife Warden, Gujarat State (WLP/RTC/28/C/57-59/2023-24).

Table 5

Reported literature on the application of ATR-FTIR spectroscopy combined with chemometric models on animal hair samples

Author	Species	Chemometric model	External validation	Result
Espinoza et al. (2009)	Asian elephant, African elephant, giraffe	Discriminant analysis	Indigenous hair artifacts (bracelets, earrings)	91.8% accuracy
Manheim et al. (2016)	Human, cat, dog	PLS-DA	Yes, another few from the same species	100% discrimination
Xu et al. (2022)	6 animal fur types	SIMCA, PCA-LDA, PLS-DA, LS-SVM	–	PLS-DA & LS-SVM: 100% accuracy; PCA-LDA: 98.33%; SIMCA: 99.44%
Prakash Sharma et al. (2022)	Pantholops hodgsonii	PLS-DA	Angora rabbit	100% discrimination; R-square value 0.99
Guleria et al. (2023)	Feather samples of 12 species of peacock	PCA-LDA	Yes, one from each of the species	100% discrimination
Bhatia et al. (2024)	Royal Bengal tiger, Indian leopard and snow leopard	PLS-DA	Human hair	100% discrimination; R-square value 0.99

Availability of data and materials

The data that support the findings of this study are not openly available due to reasons of sensitivity and are available from the corresponding author upon reasonable request.

References

- Manheim J, Doty KC, McLaughlin G, Lednev IK. Forensic hair differentiation using attenuated total reflection Fourier transform infrared (ATR FT-IR) spectroscopy. *Appl Spectrosc*. 2016 Jul;70(7):1109-17. 10.1177/0003702816652321asp.sagepub.com
- Dehury S, Mishra UK, Hembram A, Sahoo N. Light and scanning electron microscopy analysis of hair samples of some wild animals for individual identification. *J Entomol Zool Stud*. 2019;7(6):900-4.
- Ogden R, Linacre A. Wildlife forensic science: a review of genetic geographic origin assignment. *Forensic Sci Int Genet*. 2015 Sep 1;18:152-9. <https://doi.org/10.1016/j.fsigen.2015.02.008>
- Rana AK, Kumar N. Current wildlife crime (Indian scenario): major challenges and prevention approaches. *Biodivers Conserv*. 2023 Apr;32(5):1473-91.10.1007/s10531-023-02577-z
- Singh S. Over 1,000 pangolins poached and trafficked in India between 2018 and 2022. *The Hindu*. 2023 February 17. Retrieved from <https://www.thehindu.com/scitech/energy-and-environment/over-1000-pangolins-trafficked-in-india-in-5-years/article66520621.ece#:~:text=They%20are%20poached%20mainly%20for,he%20was%20quoted%20as%20saying>
- Watson K. The top five animals that poachers are hunting into extinction. *Global citizen*. 2016. Retrieved May 20, 2023, from <https://www.globalcitizen.org/en/content/most-poached-animals-worldenvironment-day/>
- Ahmed YA, Ali S, Ghallab A. Hair histology as a tool for forensic identification of some domestic animal species. *EXCLI Journal*. 2018;17:663. 10.17179/excli2018-1478
- Vaishnav L, Parashar S, Kumar A, Sachdeva MP. A study on hair analysis of different Canidae breeds. *Forensic Sci Int Rep*. 2021 Jul 1;3:100169. <https://doi.org/10.1016/j.fsir.2020.100169>
- Knecht L. The use of hair morphology in the identification of mammals. In: Huffman JE, Wallace JR, editors. *Wildlife forensics: methods and applications*. Chichester: John Wiley & Sons; 2011. p. 129-143. <https://doi.org/10.1002/9781119953142.ch8>
- Rajaselin AA, Veeraraghavan VP, Arthanari AA, Gayathri R, Kavitha S, Selvaraj J, Dinesh Y. A comparative study of different animal hairs: A microscopic analysis. *Journal of Advanced Pharmaceutical Technology & Research*. 2022 Nov;13(Suppl 1):S117.10.4103/japtr.japtr_256_22
- Farag MR, Ghoniem MH, Abou-Hadeed AH, Dhama K. Forensic identification of some wild animal hair using light and scanning electron microscopy. *Adv Anim Vet Sci*. 2015;3(10):559-68. <http://dx.doi.org/10.14737/journal.aavs/2015/3.10.559.568>
- Gharu J, Trivedi S. Hair cuticle scale patterns in hanuman langur (*Semnopithecus entellus*) and grey slender loris (*Loris lydekkerianus*). *Biological Forum – An International Journal*. 2013;5:11-15.
- Taupin JM. Forensic hair morphology comparison – a dying art or junk science?. *Sci Justice*. 2004 Apr 1;44(2):95-100. 10.1016/s1355-0306(04)71695-0
- Oien CT. Forensic hair comparison: background information for interpretation. 2009. https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/april2009/review/2009_04_review02.htm (accessed 15 May 2023).
- Sharma CP, Sharma S, Rawat GS, Singh R. Rapid and non-destructive differentiation of Shahtoosh from Pashmina/Cashmere wool using ATR FT-IR spectroscopy. *Sci Justice*. 2022 May 1;62(3):349-57. <https://doi.org/10.1016/j.scijus.2022.04.002>
- Espinoza EO, Baker BW, Moores TD, Voin D. Forensic identification of elephant and giraffe hair artifacts using HATR FTIR spectroscopy and discriminant analysis. *Endanger Species Res*. 2008 Sep 23;9(3):239-46. <https://doi.org/10.3354/esr00125>
- Xu W, Xia J, Min S, Xiong Y. Fourier transform infrared spectroscopy and chemometrics for the discrimination of animal fur types. *Spectrochim Acta A Mol Biomol Spectrosc*. 2022 Jun 5;274:121034. <https://doi.org/10.1016/j.saa.2022.121034>
- Gao H. Comparative study on the infrared spectra of straight guard hairs of leopard cat and golden cat. *Agricultural Science and Technology*. 2013 Nov 1;14(11):1683.
- Guleria P, Thakar MK, Goyal SP. Differentiation of galliformes feather samples for species identification using ATR-FTIR spectroscopy and chemometric techniques: application in wildlife forensics. *Forensic Sci Int: Animals and Environments*. 2023 Jun 9:100066. <https://doi.org/10.1016/j.fsiae.2023.100066>
- Sharma A, Chauhan R, Kumar R, Mankotia P, Verma R, Sharma V. A rapid and non-destructive ATR-FTIR spectroscopy method supported by chemometrics for discriminating between facial creams and the classification into herbal and non-herbal brands. *Spectrochim Acta A Mol Biomol Spectrosc*. 2021 Sep 5;258:119803. <https://doi.org/10.1016/j.saa.2021.119803>
- Gautam R, Vanga S, Ariese F, Umopathy S. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Techniques and Instrumentation*. 2015 Dec;2:1-38. 10.1140/epjti/s40485-015-0018-6

22. Lee LC, Liong CY, Jemain AA. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst*. 2018;143(15):3526-39. <https://doi.org/10.1039/C8AN00599K>
23. Arora T, Verma R, Kumar R, Chauhan R, Kumar B, Sharma V. Chemometrics based ATR-FTIR spectroscopy method for rapid and non-destructive discrimination between eyeliner and mascara traces. *Microchem J*. 2021 May 1;164:106080. <https://doi.org/10.1016/j.microc.2021.106080>
24. Pawani P. What is confusion matrix, accuracy, sensitivity, specificity, precision, recall? 2020. <https://poojapawani.medium.com/what-is-confusion-matrix-accuracy-sensitivity-specificity-precision-recall-1091b4723714> (accessed 15 May 2023).
25. Bhatia D, Sharma CP, Sharma S, Singh R. Species discrimination from hair using ATR-FTIR spectroscopy: Application in wildlife forensics. *Sci Justice*. 2024 May 1;64(3):314-21. <https://doi.org/10.1016/j.scijus.2024.04.002>
26. Asri MN, Verma R, Ibrahim MH, Nor NA, Sharma V, Ismail D. On the discrimination between facial creams of different brands using Raman spectroscopy and partial least squares discriminant analysis for forensic application. *Sci Justice*. 2021 Nov 1;61(6):687-96. <https://doi.org/10.1016/j.scijus.2021.08.006>
27. Asri MN, Verma R, Mahat NA, Nor NA, Desa WN, Ismail D. Raman spectroscopy with self-organizing feature maps and partial least squares discriminant analysis for discrimination and source correspondence of red gel ink pens. *Microchem J*. 2022 Apr 1;175:107170. <https://doi.org/10.1016/j.microc.2021.107170>
28. Rohart F, Gautier B, Singh A, Lê Cao KA. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol*. 2017 Nov 3;13(11):e1005752. <https://doi.org/10.1371/journal.pcbi.1005752>

ORCIDSaptarshi Rao  0000-0002-8294-6435**Corresponding author**

M.Sc. Saptarshi Rao
School of Behavioral Sciences and Forensic
Investigations
Rashtriya Raksha University
Lavad-Dahegam, Gandhinagar
IN 382305 Gujarat
e-mail: raosaptarshi@gmail.com

SZYBKA I NIENISZCZĄCA IDENTYFIKACJA SIERŚCI ZWIERZĘCEJ ZA POMOCĄ ATR-FTIR I CHEMOMETRII: MODEL KONCEPCYJNY DO STOSOWANIA W DOCHODZENIU KRYMINALNYM W PRZESTĘPSTWACH DOTYCZĄCYCH DZIKIEJ FAUNY

1. Wstęp

Włókna sierści występują wszędzie. Stanowią nieodłączny element dowodowy w przestępstwach związanych z dziko żyjącymi ssakami. Do głównych zadań prowadzących śledztwo należy ustalenie gatunku zwierzęcia na podstawie pobranej próbki sierści [1]. Analiza sierści jest techniką potencjalnie przydatną w biologii i śledztwach w sprawach przestępstw dotyczących dzikich zwierząt, umożliwiającą umiejscowienie ich nielegalnego handlu, uboju i połowu, zwłaszcza gatunków zagrożonych. Ponieważ sierść jest bardzo odporna na rozkład, nawet stare jej próbki pozwalają na zidentyfikowanie gatunku [2]. Wyróżniające poszczególne gatunki cechy sierści, takie jak kolor, tekstura, wzór osłonki, współczynnik rdzenia czy pigmentacja, pozwalają na identyfikację danego gatunku z użyciem analizy mikroskopowej. Z kolei analiza chemiczna dostarcza informacji o diecie, warunkach środowiskowych, doświadczanym przez zwierzę stresie czy jego pochodzeniu geograficznym. Wszystkie te dane mogą okazać się kluczowe do zrozumienia tendencji w przemieszczaniu się, określenia preferowanego siedliska oraz do dostarczenia informacji o nielegalnych szlakach handlowych związanych z dziką fauną [3].

Raport „Narastające przestępstwa środowiskowe”, przygotowany przez INTERPOL-UNEP, szacuje, że roczna wartość nielegalnego handlu dzikimi zwierzętami może osiągać 20 miliardów dolarów (Wildlife Crime, 2022). Biuro ONZ ds. Narkotyków i Przestępczości (United Nations Office on Drugs and Crime, UNODC) wycenia tę wartość na 7–23 miliardów dolarów. Jest to zatem czwarty pod względem skali typ nielegalnego handlu po handlu bronią, narkotykami i ludźmi. Na świecie żyje obecnie tylko niecałe 30 tys. osobników wszystkich gatunków nosorożca (w tym ponad 18 tys. osobników afrykańskiego nosorożca białego), z czego w lasach Indii żyje ich blisko 3600. Liczebność nosorożców spada drastycznie od 25 lat. Równie drastycznie spadła w ciągu ostatniego stulecia liczebność słońi: z 10 milionów do jedynie 500 tys., z czego 30 tys. osobników żyje w Indiach (a połowa z nich zamieszkuje dżungle Karnataki, Asamu i Kerali) [4]. Inny raport, sporządzony przez TRAFFIC, podaje, że w latach 2018–2022 ponad 1000 łuskowców stało się ofiarami kłusowników. W trakcie konfiskat w 50% z nich znaleziono żywe łuskowce, a w 40% – ich łuski [5]. W pułapki przeznaczone dla innych zwierząt często wpadają goryle, w następstwie czego giną

albo są chwywane i zabijane na *bushmeat* (mięso dzikich zwierząt), sprzedawane po 5–6 dolarów amerykańskich za sztukę. Według doniesień naukowych w następstwie działalności człowieka, utraty siedlisk i kłusownictwa populacja goryli zmniejszyła się do zaledwie ok. 600 osobników. Nadmierny odłów i nielegalny handel odpowiadają także za śmierć dziesiątek tysięcy żółwi każdego roku. Kilka społeczności poluje na żółwie, żeby pozyskać ich mięso, skórę i skorupy [6].

Sierść zwierzęca wielokrotnie była przedmiotem badań mikroskopowych [2, 7, 8, 9, 10, 11, 12]. Pojawiają się jednak wątpliwości związane z obiektywnością badań prowadzonych wyłącznie metodą mikroskopową, bez podłoża statystycznego [13]. Dwóch ekspertów może mieć odmienne zdanie bądź używać innej terminologii do opisu badanych cech [14]. Sierść dwóch gatunków będących przedmiotem niniejszego badania różni się kolorem. W praktyce jednak sierść najczęściej znajduje się w stanie mocno zdegradowanym, co utrudnia identyfikację. Zasada Locarda mówi, że pomiędzy stykającymi się ze sobą obiektami zachodzi wzajemna wymiana substancji. Sierść znaleziona w związku z nielegalnym zastawianiem pułapek lub w miejscach, w których doszło do użycia przemocy wobec zwierząt, może stanowić dowód w sprawie.

Obecnie najbardziej pożądaną metodą jest identyfikacja gatunków na podstawie DNA. O ile każdy gatunek ma niepowtarzalne DNA umożliwiające identyfikację, o tyle należy pamiętać, że próbki sierści często zawierają za mało tkanki cebulek włosów, co utrudnia analizę DNA [15]. W tym kontekście metoda pozwalająca na szybką, wiarygodną i nieniszczącą analizę próbek sierści jako uzupełnienie badań mikroskopowych przy określonym poziomie pewności statystycznej przyniosłaby znaczne korzyści dla dochodzeń kryminalistycznych.

Spektroskopia z transformacją Fouriera osłabionego całkowitego odbicia w podczerwieni (*attenuated total reflection Fourier transform-infrared*, ATR-FTIR) ma wiele zalet, rośnie też jej popularność w analityce i biomedycynie ze względu na jej niezawodność, szybkość i nieniszczący charakter [15]. Technika ATR-FTIR szeroko wykorzystuje się w materiałoznawstwie, chemii, biologii i kryminalistyce. W dochodzeniach przestępstw dotyczących dzikiej fauny spektroskopia ATR-FTIR w połączeniu z chemometrią stanowi skuteczne narzędzie do identyfikacji produktów pochodzących z nielegalnego handlu zwierzętami, takich jak rogi, skorupy

i korale, na podstawie ich składu chemicznego. Bardzo niewiele publikacji wyjaśnia, jak odróżnić od siebie widma różnych typów sierści w podczerwieni za pomocą metod chemometrycznych. Espinoza i in. badali użyteczność horyzontalnej ATR-FTIR w analizie i identyfikacji sierści z ogonów słoni i żyraf, z której często wyrabia się miejscową biżuterię (bransoletki, pierścionki, kolczyki i in.), w kontekście handlu dzikimi zwierzętami [16]. Inna grupa badaczy podjęła próbę wizualnego odróżnienia syntetycznego włókna od ludzkich włosów, a następnie ludzkich włosów od sierści zwierzęcej za pomocą chemometrycznych modeli danych uzyskanych techniką R-FTIR [1]. Model PLS-DA pozwolił na stuprocentowo celne odróżnianie włosów ludzkich od sierści zwierzęcej. Xu i in. porównali skuteczność czterech narzędzi chemometrycznych, tj. SIMCA, PCA-LDA, PLS-DA i LS-SVM, w analizie sześciu różnych typów futra zwierzęcego. Spośród tych czterech narzędzi PLS-DA i LS-SVM okazały się stuprocentowo skuteczne [17]. Gao za pomocą spektroskopii FTIR analizował proste włosy przewodnie u kota bengalskiego (*Prionailurus bengalensis*) i kota złotego azjatyckiego (*Catopuma temminckii*) [18]. Wyraźny pik zaobserwowano na częstotliwości 648 i 654 cm^{-1} w próbkach sierści kota złotego, ale nie kota bengalskiego. Druga pochodna widma włosa kota złotego wykazała pik o dużej intensywności, a widmo włosa lamparta – pik o małej intensywności. W ostatnim badaniu przeprowadzonym przez Sharmę i in. [15] przedstawiono innowacyjną technikę odróżniania wełny *shahtoosh* od wełny *pashmina*. Guleria i in. [19] uzyskali niedawno stuprocentową skuteczność identyfikacji próbek piór pochodzących od 12 gatunków pawia (*Galliformes*) za pomocą PCA-LDA.

Tradycyjne metody, choć powszechnie stosowane do identyfikacji różnych gatunków, mają tę wadę, że niszczą próbki [15]. Tym samym spektroskopia ATR-FTIR stanowi bardzo elastyczne z punktu widzenia kryminalistykę uzupełnienie analizy mikroskopowej badań DNA i innych stosowanych obecnie technik. Zgodnie ze stanem wiedzy autora nie opublikowano jak dotąd istotnych badań analizujących sierść różnych gatunków dzikiej fauny za pomocą ATR-FTIR. Wcześniejsze analizy dotyczyły gatunków udomowionych, które są mało istotne w badaniach dzikiej fauny. Dlatego celem niniejszego badania wstępnego było rozróżnienie widm sierści dwóch odlegle spokrewnionych gatunków – antylopy indyjskiej (*Antilope cervicapra*) i hulmana czczonego (*Semnopithecus entellus*) – za pomocą spektroskopii ATR-FTIR w połączeniu z odpowiednią techniką chemometryczną, tj. PCA i PLS-DA. Intencją autorów było pokazanie, że choć sierść tych dwóch gatunków jest podobna pod względem morfologicznym i widmowym, można ją odróżnić za pomocą modeli chemometrycznych. Niniejsze badanie może posłużyć jako prosty poradnik, przydatny do wykorzystania ATR-FTIR do badania nielegalnych

artefaktów pochodzących z dzikiej fauny, takich jak sierść, krew, kości czy zęby.

Głównym powodem stosowania chemometrii i wieloczynnikowej analizy danych ze spektroskopią ATR-FTIR jest ułatwienie interpretacji przez zmniejszenie ilości danych.

2. Materiały i metody

2.1. Pobór próbek sierści

Wzorcowe próbki sierści zostały pobrane z rejonu grzbietowego antylopy indyjskiej i hulmana czczonego w rezerwacie przyrody Indroda w dystrykcie Gandhinagar (23°14'15.2160"N, 72°38'52.0116"E). Zgodę na badania uzyskano od Głównego Konserwatora Lasów i Głównego Konserwatora Dzikiej Przyrody stanu Gujarat (WLP/RTC/28/C/57-59/2023-24). Szczegółowe informacje przedstawia tabela 1.

2.2. Przygotowanie próbek i pomiar widma

Poszczególne próbki sierści poddano sonikacji w wodzie przez 10 minut, a następnie, żeby zmniejszyć ryzyko zanieczyszczenia, przemywano alkoholem izopropylowym przez 10 minut [16]. Włosy umieszczano pojedynczo na kryształach pomiarowych wbudowanych w przystawkę ATR urządzenia. Następnie każdy włos dociskano delikatnie w połowie długości kowadełkiem dociskowym ATR w celu zapewnienia wystarczającego styku pomiędzy próbką sierści a powierzchnią kryształu przed pomiarem widma. Po zanalizowaniu danej próbki kryształ ATR każdorazowo dokładnie przecierano etanolem, a widmo tła pustej komórki ATR mierzono w tych samych warunkach. Odtwarzalność sprawdzano, analizując trzy duplikaty każdego pobranego włosa. W celu ustalenia homogeniczności rozkładu składników próbek wszystkich włosów przeprowadzono test homogeniczności polegający na badaniu każdej próbki z trzech różnych punktów (bliższego, środkowego i dalszego). Prawidłowe dane można było jednak również uzyskać na podstawie pojedynczego włosa, o ile jego średnica wypełniała co najmniej 25% okienka pomiarowego (średnica włosa 0,5 mm), przy czym najlepsze widma uzyskiwano, gdy włos całkowicie pokrywał okienko pomiarowe (2,0 mm) [16].

2.3. Oprzyrządowanie

Próbki badano za pomocą spektrometru w podczerwieni z transformacją Fouriera i diamentową przystawką Platinum do osłabionego całkowitego odbicia (Bruker alpha II ATR; Billerica, MA) zawierającego płytkę krystaliczną ZnSe w zakresie średniej podczerwieni

(MIR), tj. 4000–600 cm^{-1} przy rozdzielczości 4 cm^{-1} . Łącznie przeprowadzono 24 skanowania dla każdego widma. Dane widmowe uzyskano za pomocą programu Opus (wersja 8.1). Zastosowano detektor z deuterowanym siarczanem triglicynu i lantanem domieszkowanym α -alaniną (*deuterated lanthanum α -alanine doped triglycine sulphate detector*, RT-DLaTGS).

2.4. Przetwarzanie wstępne surowych danych matrycy

Głównym celem wstępnego przetwarzania danych było zmniejszenie szumu lub interferencji w poszczególnych widmach. Pozwoliło to skupić się na zmienności wynikającej ze składu chemicznego próbek, a nie pochodzącej ze źródeł zewnętrznych. W opisywanym tutaj badaniu zastosowano korekcję linii podstawowej do surowych widm, a dane znormalizowano przy użyciu transformacji SNV (*standard normal variate*, standardowej zmiennej normalnej), która eliminuje z widma zmienność wynikającą z nachylenia [20]. Transformacja SNV polega na zmniejszeniu wartości średniej i skalowaniu odchylenia.

2.5. Zastosowane modele chemometryczne

Chemometria jest dziedziną chemii analitycznej łączącą techniki statystyczne i matematyczne, w której uzyskuje się informacje z danych chemicznych. Opracowano modele analizy głównych składowych (*principal component analysis*, PCA) oraz analizy dyskryminacyjnej metodą częściowych najmniejszych kwadratów (*partial least squares discriminant analysis*, PLS-DA), aby rozróżnić łącznie 50 widm (25 włosów z każdego gatunku) uzyskanych od antylopy indyjskiej i hulmana czczonego w zakresie średniej podczerwieni, tj. 4000–600 cm^{-1} . Transformacji SNV dokonano w programie Unscrambler X 10.4 (CAMO Software AS, Oslo, Norwegia). Do PCA i PLS-DA użyto R-software uruchamianego w środowisku programistycznym RStudio. Zastosowano też pakiet mixOmics, opracowany specjalnie do klasyfikowania dużych ilości danych.

2.5.1. Analiza głównych składowych (PCA)

PCA zmniejsza wymiarowość danych na zasadzie redukcji złożonych zestawów danych do nowego zbioru zmiennych zwanych składowymi głównymi, które pozwalają dokładnie uchwycić cały zakres wariancji w analizowanych danych. Nienadzorowane narzędzia do redukcji wymiarowości lub modelowania dwuliniowego, takie jak PCA, są dobrze znane jako pomocne w przedstawianiu ważnych danych w bardziej zrozumiałym sposób. Składowe główne to zmienne pozwalające ustalić związek pomiędzy zmiennymi silnie skorelowanymi ze

sobą. Technika PCA wykorzystuje metodę przestrzeni własnej, czasami zwaną transformacją Karhuna-Loeve'a, i jest uznawana za kluczowe narzędzie do wizualizacji ważnych korelacji w wybranych próbkach. Wykres wag to najpopularniejszy sposób graficznego przedstawiania wyników PCA. Uwidacznia kluczowe różnice i podobieństwa pomiędzy zestawami danych, ułatwiając w ten sposób interpretację uzyskanych wyników.

2.5.2. Analiza dyskryminacyjna metodą częściowych najmniejszych kwadratów (PLS-DA)

PLS-DA to dobrze znana technika klasyfikacji liniowej łącząca specyficzne cechy regresji metodą częściowych najmniejszych kwadratów (*partial least square regression*, PLSR) ze zdolnością dyskryminacyjną metody klasyfikacyjnej. Wszechstronne podejście PLS-DA pozwala znaleźć zmienne ukryte o największej kowariancji ze zmiennymi Y i porównać je z wartościami zastępczymi zbioru danych, opierając się na metodologii kalibracji PLSR. Jeśli jakaś wartość zastępcza okaże się zgodna z wynikami, wówczas daną próbkę klasyfikuje się do tej samej kategorii, a jeśli nie – to do oddzielnej grupy klasyfikacyjnej [22, 26, 27, 28]. PLS-DA daje wiele wyników, ponieważ można ją wykorzystać jednocześnie do wyboru zmiennej dyskryminacyjnej, modelowania opisowego oraz modelowania predykcyjnego.

2.5.3. Walidacja

W celu walidacji metody i sprawdzenia skuteczności modelu zbadano po sześć niezidentyfikowanych próbek pobranych z obu gatunków. Prowadzący analizę nie miał informacji na temat tych próbek, a samych próbek nie włączono do danych treningowych. Sprawdzano, czy próbki zostały zakwalifikowane do właściwej kategorii.

3. Wyniki i omówienie

3.1. Analiza widmowa danych treningowych

Na rycinie 1 przedstawiono naniesione widma sierści pochodzącej od obu gatunków. Pod względem chemicznym sierść składa się w 65–95% z białek, przez co na jej widmie dominują piki związane z keratynami [15].

Tabela 2 zawiera listę różnych pików i odpowiadających im częstotliwości.

Jak wynika z ryciny 1, uzyskane widma mają podobny zakres daktyloskopowy z wyjątkiem jednej istotnej różnicy znajdującej się w zakresie 3400–2800 cm^{-1} . Tylko jeden wyraźny pik na 3289 cm^{-1} w sierści mały może wynikać z rozciągania grupy O-H powiązanej z H i symetrycznym rozciąganiem N-H (woda i grupa hydroksylowa). Natomiast w widmie sierści antylopy pojawiają

się dwa piki na częstotliwościach 2921 i 2851 cm^{-1} odpowiadające rozciąganiu C-H o średniej intensywności. Zaobserwowano też dwa wyraźne piki w widmie sierści obu gatunków na częstotliwościach 1634 i 1516 cm^{-1} , związane odpowiednio z amidem I i II. Pozostałe piki uznano za słabe z powodu ich małej intensywności. Pik na ok. 1454 cm^{-1} wynika zarówno z antysymetrycznej deformacji grup CH_3 , jak i trybu nożycowego grup CH_2 , podczas gdy pik przy 1394 cm^{-1} był spowodowany grupami aminokwasowymi lub deformacją grup metylowych.

Widma najpierw analizowano wzrokowo, a następnie technikami PCA i PLS-DA. Przeprowadzono test homogeniczności na wszystkich próbkach sierści polegający na analizowaniu każdego włosa w trzech różnych pozycjach (lewej, środkowej i prawej), a odtwarzalność sprawdzano, testując trzy duplikaty każdego włosa. W przeciwieństwie do oczekiwań bezpośrednie badanie wzrokowe nie ujawniło żadnej zmiany w nakładanych widmach, co pokazano na rycinie 1. Oba widma są prawie takie same; bardzo niewielkie są też różnice w wartościach długości fali, mimo że badane gatunki są daleko spokrewnione. Odróżnienie ich widm wyłącznie na podstawie inspekcji wzrokowej jest dość trudne. Inspekcja taka jest z natury subiektywna, stąd potrzeba zastosowania modeli chemometrycznych w celu odróżnienia gatunków i uzyskania niezbędnych informacji.

3.2. Analiza głównych składowych

Nieprzetworzone dane wprowadzono do R-software w celu analizy PCA. Aby zapewnić wiarygodność wyników, obie kategorie próbek sierści były równej wielkości, tj. zawierały po 25 próbek sierści antylopy i mały. Na rycinie 2 przedstawiono dwuwymiarowy wykres próbek sierści, częściowo uwidaczniający ich podział. Główna składowa 1 (PC1) wyjaśniała 62% wariancji w zbiorze danych, a główna składowa 2 (PC2) – 25%. Razem za pomocą PC1 i PC2 osiągnięto łączną wariancję na poziomie 87%. Stworzono ponadto wykres osypiska w celu ustalenia, ile głównych składowych pozwoliłoby na osiągnięcie największej wariancji. Jak wynika z ryciny 2, PCA dostarcza informacji ilościowych na temat liczby klas, ale nie dzieli próbek na różne klasy.

Ponieważ PCA jest techniką nienadzorowaną, przeznaczoną do wstępnych badań, przeprowadzono PLS-DA w celu uzyskania lepszej dyskryminacji i przewidywania nieznanych próbek.

3.3. Rozróżnianie za pomocą modelu PLS-DA

W przeciwieństwie do inspekcji wzrokowej i PCA, celem PLS-DA było ustalenie wiarygodnej, zadowalającej kategoryzacji dwóch badanych typów sierści. Model PLS-DA skonstruowano na podstawie treningowego zestawu 50 próbek. W celu rozróżnienia i kategoryzacji

próbek sierści antylopy indyjskiej i hulmana czczonego otrzymany model zawierał dwie wymagane zmienne ukryte. Podejście PLS-DA zastosowano do danych znormalizowanych w R za pomocą pakietu mixOmics, który umożliwia sprawne opracowywanie dużych zestawów danych i konfigurowalne użycie próbek [23]. Na rycinie 3 przedstawiono dwuwymiarowy wykres podziału na klasy za pomocą PLS-DA. Oba typy włosów zostały błędnie skategoryzowane do odpowiadających im grup, co widać na obu rycinach. Jak pokazuje rycina 4 (lewa strona), po wybraniu 10 komponentów obliczono, że błąd zbalansowany wyniósł 0. Krzywą charakterystyki operacyjnej odbiornika (ROC) dla dwóch komponentów pokazuje rycina 4 (prawa strona), wraz z wartością AUC wybitnego modelu klasyfikacji, wynoszącą 1.

3.4. Walidacja

Rycina 5 przedstawia podział niezidentyfikowanych próbek, oznaczonych czarnymi kropkami, na odpowiednie klasy. Nieznane próbki oraz ich przewidywane klasy przedstawiono w tabeli 3.

Jak pokazuje tabela 4, sporządzono również macierz błędów, aby sprawdzić czułość i specyficzność modelu. Liczbę prawdziwie dodatnich prognoz dzieli się przez sumę liczby prawdziwie dodatnich i fałszywie ujemnych prognoz w celu określenia czułości [24]. W tym przypadku czułość wyniosła 1,0. Specyficzność oblicza się, dzieląc sumę liczby wskazań fałszywie dodatnich i prawdziwie ujemnych przez liczbę wskazań prawdziwie ujemnych. W tym przypadku specyficzność również wyniosła 1,0, podobnie jak dokładność określona na podstawie macierzy błędów (tj. osiągnięto 100% dokładności).

$$\text{Czułość} = \text{PD}/(\text{PD} + \text{FU}) \times 100$$

$$\text{Specyficzność} = \text{PU}/(\text{PU} + \text{FD}) \times 100$$

$$\text{Dokładność} = \text{PD} + \text{PU}/(\text{PD} + \text{PU} + \text{FD} + \text{FU}) \times 100$$

Rozróżnianie i identyfikacja gatunków to bardzo ważne elementy mające wiele zastosowań w kryminalistyce. W opisywanym tutaj badaniu za pomocą spektroskopii ATR-FTIR i chemometrii rozróżniano sierść dwóch daleko spokrewnionych gatunków dzikiej fauny. Podjęto próbę ich identyfikacji i rozróżnienia na podstawie zaproponowanej techniki. Ponieważ inspekcja wzrokowa była powolna i wiązała się z ryzykiem błędów, konieczne było sięgnięcie po narzędzia chemometryczne, które usprawniły analizę. Łączna wariancja metody PCA wyniosła 87%. Metoda ta pozwoliła do pewnego stopnia rozróżnić próbki, dostarczyła również o nich informacji ilościowych. Z kolei PLS-DA rozróżniła obie kategorie próbek ze stuprocentową dokładnością. Stwierdzono, że zarówno czułość, jak i specyficzność modelu wynosi 1,0, co wskazuje, że PLS-DA to wyjątkowo skuteczne narzędzie do rozróżniania próbek. Przeprowadzono ślepe próby na sześciu nieznanymi próbkach za pomocą

opracowanego modelu PLS-DA. Wszystkie te próbki zostały bezbłędnie przydzielone do odpowiednich grup. Tabela 5 przedstawia wszystkie przeprowadzone badania związane z tą tematyką.

4. Wnioski

Opisane w niniejszej pracy badanie jest propozycją wstępnej – ze względu na niewielką próbę (tylko dwa gatunki) – metody rozróżniania sierści; dostarcza także perspektywnego narzędzia, przydatnego w tego typu rozróżnianiu. Autorzy chcą ponadto zaznaczyć, że proponowana metoda jest nie tyle zamiennikiem analizy DNA, co jej uzupełnieniem. Z powodów etycznych wielkość próby była mała. Do zbadania pełnych możliwości spektroskopii ATR-FTIR potrzebne są większe i bardziej szczegółowe badania. Warto też zauważyć, że opisana tu metoda rozróżniania może okazać się skuteczna w przypadkach, gdy dowody w postaci włosów znalezione w podrobionych przedmiotach są mocno zniszczone, a analiza DNA nie przyniesie oczekiwanych rezultatów. Jeśli proponowana tu metoda potrafi rozróżnić i zidentyfikować gatunki daleko spokrewnione, to może także rozróżnić i zidentyfikować gatunki bardziej morfologicznie zbliżone. Innym celem niniejszego badania była dalsza analiza metody na podstawie próbek sierści różnych gatunków. Obecnie dostępne przenośne spektrometry ATR-FTIR można zastosować do badań w terenie. Oprócz sierści przedmiotem badań mogą być zęby, kości czy inne dowody pochodzenia zwierzęcego, które często spotyka się w związku z nielegalnym handlem.

Jednocześnie wnioskiem z niniejszego badania jest to, że spektroskopia ATR-FTIR i chemometria stanowią rzetelne, nieniszczące metody identyfikacji gatunków dzikiej fauny w śledztwach dotyczących przestępstw kłusownictwa na dzikich zwierzętach.

Zgoda komisji bioetycznej

Wszystkie procedury badań z udziałem zwierząt były zgodne z wytycznymi Dyrektywy Rady z dnia 24 listopada 1986 r., a zgodę na nie uzyskano od Głównego Konserwatora Lasów i Głównego Konserwatora Dzikiej Przyrody stanu Gujarat (WLP/RTC/28/C/57-59/2023-24).

Dostępność danych i materiałów

Dane potwierdzające wyniki tego badania nie są powszechnie dostępne z powodu ich wrażliwości. Dostęp do nich można uzyskać od autora korespondencyjnego na uzasadnioną prośbę.