International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 8 Issue 3, May-June 2024 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Classification of Lumpy Skin Disease using Principal Component Analysis (PCA)-based Supervised Machine Learning

Rajesh Kumar[1], Dr. Sneha Soni[2]

[1]Research Scholar, [2]Head and Professor,
[1,2]Department of CSE, SIRTE, Bhopal, Madhya Pradesh, India

ABSTRACT

A viral illness, lumpy skin in cattle is spread by mosquitoes and other insects that feed on human blood. Animals that have never been exposed to the virus are mostly affected by the sickness. Milk, meat, and domestic and international livestock commerce are all impacted by cattle lumpy skin disease. Traditional lumpy skin disease diagnosis is exceedingly time-consuming, complicated, and resource-constrained. As a consequence, it is essential to use deep learning algorithms that can categorize the condition with excellent performance outcomes. In order to segment and classify diseases using deep features, deep learning-based segmentation and classification are suggested. Convolutional neural networks with 10 layers have been selected for this. The created framework is first trained using data gathered from cattle with Cattle's Lumpy Skin Disease (CLSD). The skin tone is crucial to identifying the damaged region when a disease is represented since the characteristics are derived from the input photographs. To do this, a color histogram was utilized. A deep pre-trained CNN uses this divided region of altered skin color to extract features. Next, a threshold is used to transform the produced result into a binary format. The classifier for classification is PCA-Driven Supervised Machine Learning. The suggested methodology's classification performance has a 96% CLSD accuracy rate. We give a comparison with cutting-edge methodologies to demonstrate the efficacy of the suggested strategies.

KEYWORDS: *CLSD, CNN, PCA-Driven Supervised Machine Learning, Deep Learning, Transfer Learning*

How to cite this paper: Rajesh Kumar | Dr. Sneha Soni "Classification of Lumpy Skin Disease using Principal Component Analysis (PCA)-based Supervised Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-8 | Issue-3, June 2024, pp.648-653, URL: www.ijtsrd.com/papers/ijtsrd64954.pdf

IJTSRD64954

I. INTRODUCTION

Skin illnesses include a wide spectrum of ailments that affect the skin and may be brought on by bacterial, viral, fungal, allergy, and parasitic infections, as well as by the skin itself. Skin infections are illnesses that impact the skin. These illnesses have the potential to result in rashes, itching, inflammation, and other skin abnormalities. While some skin diseases are inherited, others are brought on by lifestyle choices. Ointments, creams, medicines, and dietary modifications are all used as treatments for various skin disorders.

The human body is covered by and protected by the skin, a sizable organ. It performs a number of different tasks, including:

- Holding in liquids and avoiding dehydration
- Preventing the spread of germs, viruses, and other disorders
- Assisting you in experiencing various feelings, such as pain or temperature
- Producing vitamin D;
- Controlling body temperature

All disorders that irritate, swell, and irritate the skin and cause rashes and other changes to the look of the skin are considered skin issues. A total of 1.79% of all illnesses in the globe are caused by skin conditions. The American Academy of Dermatology Association states. In the US, 1 in 4 people have skin problems. Skin conditions may range widely in intensity and symptoms, and they can be either transient or permanent, painful or not. Some skin infections may not be serious, while others may be fatal.

These anomalies are crucial in the identification of a number of internal disorders, even though the majority of skin diseases start in the layers of the skin. The idea that a person's skin reflects their inside health has some validity. The skin is often the first organ of the body to exhibit observable symptoms of underlying illness because of its visibility and accessibility. Skin anomalies typically point to metabolic, cancerous, and glandular illnesses.

Skin is subject to pathological alterations of various kinds, including genetic, inflammatory, benign and malignant (neoplastic), endocrine, hormonal, traumatic, and degenerative processes, much like other tissues. The condition of the skin is also impacted by emotions. In many respects, the skin's response to various illnesses and disorders is different from that of other tissues. Extensive skin inflammation, for instance, may interfere with the metabolism of other organs and bodily systems, resulting in anemia, circulatory collapse, temperature abnormalities, and disturbances of the water and electrolyte balance in the blood. However, due to the skin's powerful healing abilities, extensive harm, such as thermal burns, may be followed by a significant amount of regrowth of the diseased or wounded regions, with a disproportionally low degree of scarring.

II. BACKGROUND

LSD is a serious transboundary sickness that affects the steers business globally, as indicated by Ayesha Anwar et al. [2]. The point of this study was to identify designs and huge change focuses as well as to foresee the recurrence of LSD pandemic reports in Africa, Europe, and Asia. Information from the LSD pandemic report from the World Association for Creature Wellbeing, which traversed the years January 2005 through January 2022, were investigated. Auto-backward moving normal (ARIMA) and brain network auto-backward (NNAR) models were utilized to figure the quantity of LSD reports, and paired division was used to pinpoint genuinely significant information change focuses. Four crucial defining moments were recognized, one for every mainland. The period between the third and fourth change focuses (2016-2019) in the African information had the most noteworthy mean number of LSD reports. Gigantic episodes happened somewhere in the range of 2015 and 2017 and related with each adjustment of the LSD pandemic in Europe. Asia had the most LSD grievances in 2019 after the third archived changeover point in 2018. In the following three years (2022-2024), both ARIMA and NNAR figure an ascent in LSD reports in Africa and a consistent number in Europe. Be that as it may, in spite of the fact that ARIMA predicts a steady number of episodes in Asia in 2023-2024, NNAR conjectures an expansion in such flare-ups. This study's discoveries add to our developing information on the study of disease transmission of LSD.

Azeem as well as others [3], Ongoing worldwide extension of the uneven skin sickness infection beyond endemic sub-Saharan nations in the Center East and Asia proposes transboundary transmission. LSD pandemics have been found in Asia's Bangladesh, India, China, Nepal, Bhutan, Vietnam, Myanmar, Sri Lanka, Thailand, Malaysia, and Laos interestingly. The dairy and animals ventures in the space are seriously worried by this. This report sums up information on prior episodes in southern Asia and accentuates the danger that LSD postures to adjoining nations. The various systems and activities expected to control episodes of this as of late distinguished disorder in Asia are additionally encouraged.

Among others, episodes of Punyapornwithaya [4] and Knotty skin illness (LSD) impacted cow cultivates all over Thailand in 2021-2022. The LSD pandemic that cleared the nation started with this episode. Therefore, a more full comprehension of LSD the study of disease transmission is required. This study set planned to find the spatiotemporal examples of LSD episodes in regions with a critical number of dairy ranches. Information from LSD flare-up examinations assembled from dairy ranches in Khon Kean area, northern Thailand, were analyzed utilizing spatio-fleeting models such space-time stage, Poisson, and Bernoulli models. 133 out of 152 dairy ranches had LSD episodes among May and July 2021. Most of dairy ranches (n = 102) were altogether affected by the LSD episodes in June. The paces of horribleness and mortality from all crowd assaults were 87, 31, and 0.9%, separately. The groups in the review region that were considered to be the most probable in view of the consequences of all models were situated in its northernmost area. The Poisson and space-time change models distinguished 15 and 6 spatio-worldly flare-up bunches, separately, however the Bernoulli model just found one group. The most probable, not entirely settled by those reenactments, had radii of 1.59, 4.51, and 4.44 km, individually. All ranches that were a part of the bunch distinguished by the space-time stage model were likewise recognized by the Poisson model, showing that the pestilence zone was distinguished by the two procedures. As per the review's outcomes, ranchers who own fields inside a one-kilometer span of where the LSD flare-up began need to utilize more rigid bug vector control techniques to check the infection's spread. This study adds to a superior comprehension of the spatiotemporal example of LSD groups in the pestilence region. The discoveries of this study might help chiefs in thinking up methodologies to forestall and control future scourges as well as in focusing on the appropriation of assets to high-gamble with regions.

Notwithstanding Mishra [5], Thailand is one of the nations where foot and mouth illness flare-ups have brought about impressive monetary misfortunes. A critical early-cautioning technique that could help the public authority in laying out a FMD checking and control program is guaging. The objective of this study was to display and gauge the event of FMD flare-up episodes (n-FMD episodes) consistently in Thailand utilizing time-series techniques like the occasional autoregressive coordinated moving normal (SARIMA), blunder pattern irregularity (ETS), brain network autoregression (NNAR), geometrical remarkable smoothing state-space model with Box-Cox change, ARMA mistakes, Pattern and Occasional parts (TBATS), and half breed strategies. These systems were utilized to the month to month n-FMD episodes (n = 1209) between January 2010 and December 2020. The discoveries showed that the n-FMD events had a steady pattern from 2010 to 2020, however from 2014 to 2020, they were by all accounts turning out to be more terrible. There is an occasional pattern to the pandemic events, with the yearly pinnacle frequently happening among September and November. The single-method time-series models, like TBATS (1,0,0), SARIMA (1,0,1) (0,1,1) (12), NNAR (3,1,2)12, and ETS (A,N,A), produced the best fit. SARIMA-NNAR and NNAR-TBATS were the mixture models that played out the best on the approval datasets. Models that incorporate irregularity and a non-direct pattern perform better compared to different models in examination. The conjectures featured the rising pattern of n-FMD episodes in Thailand, which borders a few countries with an endemic FMD plague and where cross-line traffic in cows is normal. Accordingly, making compelling control techniques and FMD scourge anticipation methodologies in both Thailand and its neighbors is basic.

Perone and extra [6], The strength of cows is a significant worry for people in general. At the point when sicknesses in creatures are distinguished in the beginning phases, they may at times be dealt with. The livestock business might experience enormous monetary misfortunes in the event that uneven skin sickness isn't effectively treated. The primary driver of this disease is the uneven skin infection, which is an individual from the Poxviridae family. The significant side effect of uneven skin sickness is the Neethling strain, and different side effects incorporate a couple of gentle varieties of restricted skin knobs. These side effects additionally affect the mucous layers of inner organs, like the regenerative and respiratory frameworks. Creatures with this infection, like dairy cattle, have for all time harmed skin. Contaminated cows produce less milk, are bound to be sterile, have unfortunate development, have fetus removals, and even fade away. In this review, an engineering in light of AI is made to distinguish the affliction. After highlight extraction from pretrained models like VGG-16, VGG-19, and Beginning v3, this engineering utilizes a few classifiers. We accordingly characterized the gathered elements utilizing our physically developed dataset and classifiers including kNN, SVM, NB, ANN, and LR to test the work. With this technique, the cutting-edge arrangement had a grouping exactness of 92.5% for the test dataset.

III. FEATURE SELECTION

The process of choosing a subset of the most important characteristics to be utilized in the model-building process is known as feature selection, one of the crucial and often used machine learning methods in data mining. To maximize classification accuracy with comparatively fewer features, feature selection is therefore often thought of as an optimization issue. In order to do this, duplicate and unnecessary characteristics from raw datasets are often removed via feature selection. Since the 1970s, feature selection approaches have been extensively used in a broad range of domains, including the prediction of protein structure classes, the categorization of tracked neurons, text classification, the detection of auditory events, and the classification of gene expression data. In order to choose the marker genes for lumpy skin disease that influence the classification accuracy, feature selection approaches are also applied. The microarray-based molecular categorization of illnesses has made significant strides, but it is still a long way from being used in clinical practice. Numerous feature selection techniques, including the Fisher score, have been used to date in the choice of feature genes. One of the most popular supervised feature selection techniques is the Fisher score. The method we'll use delivers, in decreasing order, the rankings of the variables based on the fisher score. The variables may then be chosen based on the situation.

IV. METHODOLOGY

In this work, we used a transfer learning strategy to classify lumpy skin diseases using the trained supervised learning model. The output of the pre-trained model was flattened and supplied to the classifier using picture data alone for both binary and multi-class classification of LSDs. The classifier then classifies the LSD using the concatenation of both the picture data and the cattle information. A dense layer with 128 neurons was added on top of the pre-trained supervised learning model since the picture data it produces is greater than the patient information feature. As a result, the model's output image features are reduced to 128 and the classifier's two input types—one-hot encoded patient information features and image features—are balanced. In order to address the dataset's class imbalance, a weighted loss function based on label frequency was used, which gives fewer represented classes greater weight. Sure, here is the pseudo code for a supervised machine learning algorithm using PCA for dimensionality reduction:

Step 1. Data Preparation

After loading the dataset, divide it into labels ({y}) and features ({X}).

DATASET<- import the dataset

Divide the data into test and training sets.

X, y <- divide the dataset into labels and features

divide X and Y into training and test sets (e.g., 80% train, 20% test) X_train, X_test, y_train, y_test

Step 2. Applying PCA

- Set the required number of components ({n_components}) at the beginning of the PCA model.
  PCA_MODEL <- Use n_components to initialize PCA
- To get the reduced training set ({X_train_pca}), fit the PCA model on the training set ({X_train}) and transform it.
  X_train_pca \- fit X_train to PCA_MODEL and convert X_train
- To get the reduced test set ({X_test_pca}), transform the test set ({X_test}) using the fitted PCA model.
  X_test_pca \- Apply PCA_MODEL to convert X_test

Step 3. Model Training

- Set up a supervised learning model (like Logistic Regression) from scratch.
  MODEL \- Set up a supervised learning model, such as Logistic Regression.
- Use the associated labels ({y_train}) and the shortened training set ({X_train_pca}) to train the model.
  fit MODEL for both y_train and X_train_pca

Step 4. Model Evaluation

- Forecast both the test set ({X_test_pca}) and the training set ({X_train_pca}).
  y_train_pred <- forecast on X_train_pca using MODEL
  y_test_pred <- forecast on X_test_pca using MODEL
- Determine the model's accuracy using both the training and test sets.
  testing_accuracy \- use y_test and y_test_pred to determine accuracy; training_accuracy \- use y_train and y_train_pred to calculate accuracy
- Create and print the test set prediction classification report and confusion matrix.

Classification report generated by utilizing y_test and y_test_pred print CONFUSION_MATRIX {- compute confusion matrix using y_test and y_test_pred The training accuracy print is "training_accuracy:". The print "Test Accuracy:", test_accuracy The print "Confusion Matrix:", CONFUSION_MATRIX CLASSIFICATION_REPORT, "Classification Report:"

The general procedure for fusing PCA with a supervised machine learning method is described in this pseudo code. Modify the particular stages according on your dataset and needs.

V. RESULTS AND ANALYSIS

Following the pre-processing of the data in this experiment, the top ten features of the dataset are chosen using the features selection technique, and then well-known supervised and unsupervised learning methods are put into practice. Support Vector Machine (SVM), XGBoost, and Extreme Learning Machine (ELM) are used in machine learning. The suggested proposed model belongs to the Efficient CNN subcategory.

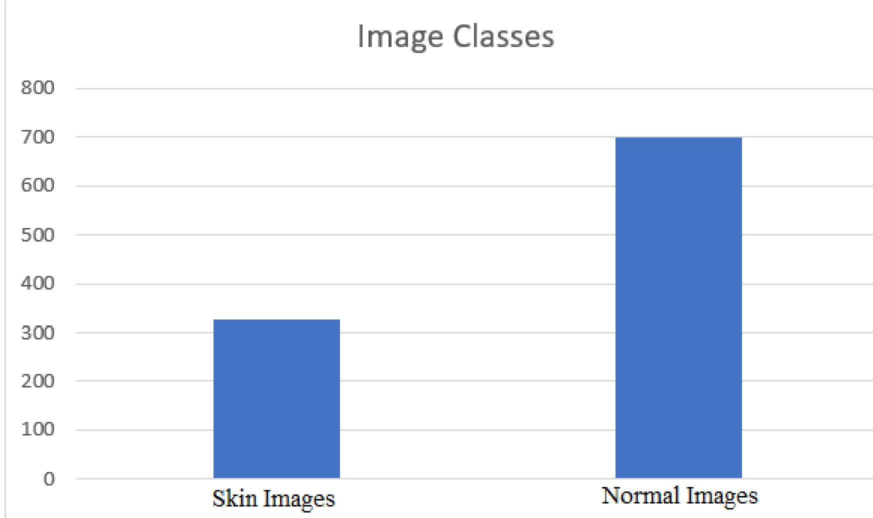

Figure 2: Instances in Dataset

Table 1: Experiment Results of Machine Learning Models

Method	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)
C-SVM	88.89	89.98	989.51	89.08
C-KNN	87.80	88.00	88.10	87.82
Q-SVM	89.01	89.20	89.00	89.03
ESD	89.00	89.89	90.00	89.80
M-SVM	89.00	89.00	90	89.11
XGBoost	90.01	89.00	89.25	89.92
ELM	90.01	90.05	90.19	90.06

Table 1 displays the experiment results for various machine learning methods, whereas Table 1 displays the experiment results for various machine learning algorithms, including the suggested model. The comparison between the suggested model and the state-of-the-art model is shown in Table 2.

Table 2: Experiment Results of Different Machine Learning Algorithms including the Proposed Model

Method	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)
ESD	89.00	89.89	90.00	89.80
M-SVM	89.00	89.00	90	89.11
XGBoost	90.01	89.00	89.25	89.92
ELM	90.01	90.05	90.19	90.06
Proposed	91.31	90.84	92.14	94.11

In this part, we compare the outcomes produced by our suggested simulation model to those produced by earlier suggested models. Table 4.3 displays our findings and contrasts them with findings from other models.

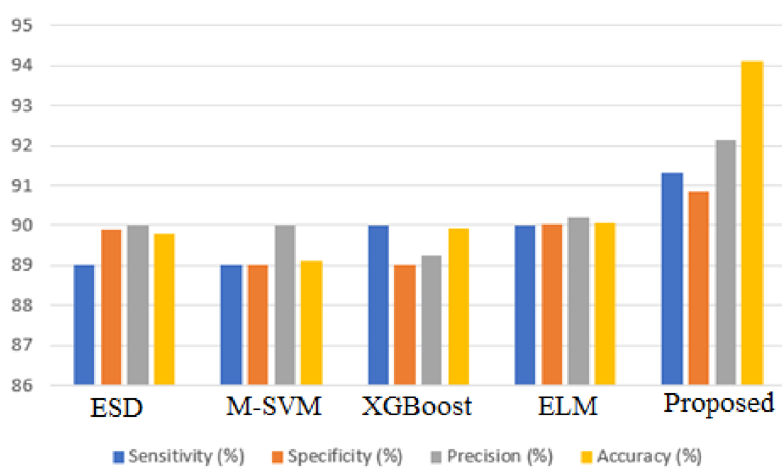

Figure 3: Performance Analysis of Different Models

VI. CONCLUSIONS

This study offered a segmentation and classification model for the lumpy skin condition of cattle. A transfer learning approach using MobiltNetV2 was detailed in the framework. On the well-known datasets for the lumpy skin disease in cattle, the suggested technique was assessed. Different kinds of supervised and unsupervised learning approaches are analyzed using F1-Measure, recall, precision, and recall, which are briefly addressed. The Fisher score approach was used to examine and pre-process the LSD after it was obtained from the Kaggle library. This method minimizes the amount of features in the dataset and prevents the over-fitting issue. On the pre-processed dataset, supervised and unsupervised machine learning methods are used. The ensemble model outperforms all other models including the state-of-the-art model when the performance of all the methods is compared.

The following are the thesis's future focuses:

- To provide a thorough analysis of Deep Learning algorithms using a real-time dataset to offer a better solution for the intrusion detection system.
- Looking at novel pre-processing methods that might increase model accuracy.
- The performance of the LSD detection system may be improved by using deep learning methods.

REFERENCES

[1] Musa Genemo, "Detecting high-risk Area for Lumpy Skin Disease in Cattle Using Deep Learning Feature", Advances in Artificial Intelligence Research (AAIR), Vol. 3 (No. 1), pp. 27-35, ISSN 2757-7422, (2023).

[2] Ayesha Anwar, Kannika Na-Lampang, Narin and Veerasak, "Lumpy Skin Disease Outbreaks in Africa, Europe, and Asia (2005–2022): Multiple Change Point Analysis and Time Series Forecast", MDPI Journal, (2022).

[3] Azeem, S.; Sharma, B.; Shabir, S.; Akbar, H.; Venter, E., "Lumpy skin disease is expanding its geographic range: A challenge for Asian livestock management and food security". Vet. J. (2022), 279, 105785.

[4] Punyapornwithaya, V.; Seesupa, S.; Phuykhamsingha, S.; Arjkumpa, O.; Sansamur, C.; Jarassaeng, C. Spatio-temporal patterns of lumpy skin disease outbreaks in dairy farms in northeastern Thailand. Front. Vet. Sci. (2022), 9, 957306.

[5] Punyapornwithaya, V.; Mishra, P.; Sansamur, C.; Pfeiffer, D.; Arjkumpa, O.; Prakotcheo, R.; Damrongwatanapokin, T.; Jam-pachaisri, K. Time-Series Analysis for the Number of Foot and Mouth Disease Outbreak Episodes in Cattle Farms in Thailand Using Data from 2010–2020. Viruses (2022), 14, 1367.

[6] Perone, G. Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy. Eur. J. Health Econ. (2022), 23, 917–940.

[7] Maw, M.T.; Khin, M.M.; Hadrill, D.; Meki, I.K.; Settypalli, T.B.K.; Kyin, M.M.; Myint, W.W.; Thein, W.Z.; Aye, O.; Palamara, E. First Report of Lumpy Skin Disease in Myanmar and Molecular Analysis of the Field Virus Isolates. Microorganisms (2022), 10, 897.

[8] Singhla, T.; Boonsri, K.; Kreausukon, K.; Modethed, W.; Pringproa, K.; Sthitmatee, N.; Punyapornwithaya, V.; Vinitchaikul, P. Molecular Characterization and Phylogenetic Analysis of Lumpy Skin Disease Virus Collected from Outbreaks in Northern Thailand in (2021). Vet. Sci. 2022, 9, 194.

[9] Khalil, M.I.; Sarker, M.F.R.; Hasib, F.Y.; Chowdhury, S. Outbreak investigation of lumpy skin disease in dairy farms at Barishal, Bangladesh. Turk. J. Agric. Food Sci. Technol. (2021), 9, 205–209.

[10] Lu, G.; Xie, J.; Luo, J.; Shao, R.; Jia, K.; Li, S. Lumpy skin disease outbreaks in China, since 3 August 2019. Transbound. Emerg. Dis. (2021), 68, 216–219.

[11] Chibssa, T.R.; Sombo, M.; Lichoti, J.K.; Adam, T.I.B.; Liu, Y.; Elraouf, Y.A.; Grabherr, R.; Settypalli, T.B.K.; Berguido, F.J.; Loitsch, A. Molecular analysis of East African lumpy skin disease viruses reveals a mixed isolate with features of both vaccine and field isolates. Microorganisms (2021), 9, 1142.

[12] Fatemeh Namazi and Azizollah Khodakaram Tafti, "Lumpy skin disease, an emerging transboundary viral disease: A review", Vet Med Sci. (2021) May; 7(3): 888–896.