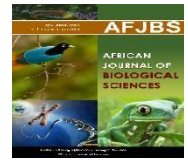
https://doi.org/10.48047/AFJBS.6.8.2024.496-519

African Journal of Biological Sciences



Research Paper **Open Access**

A COMPREHENSIVE SURVEY OF DATA PRE-PROCESSING TECHNIQUES FOR AUDIO, VIDEO, AND TEXT: APPROACHES AND APPLICATIONS

Dr. Vani H Y[1]*, Vanishri Arun[2], Anusuya M A[2], Shashikala B.M[3], Ashwini Jadhav[3], Madhusudhan G[4], Smitha M[4], Mayura D. Tapkire[5], Lavanya M. S[6]

[12346]JSS Science and Technology University Mysuru, India

[5]The National Institute of Engineering, Mysuru, India

[3]Ashwini Jadhav, Science Teaching and Learning Centre University, University of Witwatersrand Johannesburg, South Africa

Article History

Volume 6,Issue 8, 2024

Received:20 Mar 2024

Accepted : 23 Apr 2024

doi: 10.33472/AFJBS.6.8.2024.496-519

ABSTRACT

This research paper provides a comprehensive examination of data pre-processing methods specifically tailored for the analysis of extensive building operational data. It covers a wide range of pre-processing techniques within the domains of Audio, Video, and Text analysis, delving into their impacts on accuracy. The study also addresses multifaceted challenges inherent to voice recognition, emotion detection through speech and text, facial recognition, unique text identification, and image segmentation. Comparative evaluations of pre- and post-preprocessing outcomes are presented for image, text, and audio data, using standard datasets. The research spans over a decade, and its results are summarized in tabular form. Effective data preprocessing is crucial due to the intricate and uncertain nature of the examined data. The paper elucidates an array of preprocessing techniques for tasks like missing value imputation, outlier detection, data reduction, data scaling, data transformation, and data partitioning. Additionally, the study introduces advanced data science methodologies, including data augmentation, transfer learning, and semi-supervised learning, to address practical challenges in building-related data analysis. The research critically assesses the strengths and limitations of existing preprocessing methodologies, outlines potential future research directions, and envisions practical applications in the field of intelligent building energy management. This paper's novelty lies in its holistic approach to data preprocessing techniques across multiple domains, encompassing Neural Network Models in image processing, Speech Processing, Text Mining, Genetic Algorithm-Based Machine Learning Techniques, Hybrid Algorithm-Based Machine Learning Techniques, and Fuzzy-Based Machine Learning Techniques, providing a unified framework for evaluating these techniques beyond individual areas.

Keywords: *Data pre-processing, Data Augmentation, SMOTEK, Neural network, Speech Processing*

INTRODUCTION

Data analysis is a fundamental pursuit, spanning engineering, science, and process control, aiming to extract latent knowledge from diverse sources such as sensors, machinery, networks, and devices. This raw data is often complex and must be meticulously analyzed to unveil hidden trends, enabling comprehension and prediction. However, the quality of the insights derived from data analysis is inherently linked to the quality of the data itself. Data preprocessing, as a preliminary or concurrent stage of analysis, addresses issues like data corruption and missing attributes, enhancing data suitability. This paper is dedicated to several focal points within data preprocessing:

Application of Filters: Filters are employed to clean image, speech, and text data by removing unwanted patterns.

Augmentation Process: Techniques for enhancing image, speech, and text data repositories are explored.

SMOTETomek Method: Dealing with imbalanced data in image, speech, and text datasets through the application of the SMOTETomek method.

Feature Selection Methods: Optimizing feature selection in image, speech, and text data to improve model performance.

The paper begins with an overview of data preprocessing and its rationales and then delves into a wide array of data preprocessing techniques. It further elaborates on the myriad preprocessing techniques embraced by researchers and demonstrates the application of these methods across diverse datasets. In the concluding section, novel data preprocessing methodologies are introduced.What sets this study apart is its incorporation of advanced techniques, including data augmentation, convolutional neural networks (CNNs), and edge detection in image preprocessing. These techniques extract richer information from images, while speech processing is explored for emotion recognition, classification, and edge detection across datasets. Text mining is applied to various textual data sources, uncovering hidden patterns, and fuzzy inference systems are integrated with machine learning methods to enhance model performance. These unique elements distinguish this study as a comprehensive and pioneering work in data preprocessing techniques. The research has meticulously considered parameters within each domain, including image processing, speech processing, text mining, genetic algorithm-based machine learning, hybrid algorithm-based machine learning, and fuzzy-based machine learning. These parameters have been systematically analyzed to provide a robust framework for data preprocessing, applicable in real-world scenarios across diverse domains.

LITERATURE SURVEY

The rapid advancements in artificial intelligence and machine learning have propelled the development of powerful neural network models for various applications, especially in image processing. These models have demonstrated remarkable capabilities in tasks such as image classification, object detection, and facial recognition. However, their success hinges not only on the complexity of the architecture but also on the preprocessing steps that lay the foundation for effective model training and prediction. Data preprocessing [5, 6] is a pivotal stage in the machine learning pipeline, involving techniques that transform raw data into a suitable format for analysis. This literature survey focuses on a comprehensive exploration of neural network models applied to image processing, shedding light on the preprocessing methods employed to enhance their performance. Table 1 displays the outcomes of data preprocessing techniques applied to a range of datasets encompassing audio, video, speech and text data.

Author (year)	Dataset	Purpose	Methods	Result	Data Preprocessing	Remarks	
Neural Network Models-image processing							
Swagata [7] Boruah (2023)	APTOS [8] 2019	Diabetic Retinopathy (DR)	ResNet2.0 mode	91%-VA	Preprocessing steps included up-sampling, applying Gaussian blur, resizing the images, and converting them to RGB format. Additionally, data augmentation techniques were applied.	Noteworthy aspects of the preprocessing phase included the use of both negative and positive weighted Gaussian blur filters, noise up-sampling, and addressing class imbalance concerns.	Able to detect DR at the initial stages of mild and moderate DR. to reduce the complexity of the ResNet model while maintaining its ability to achieve excellent results.
Ali Bakhshi [9] (2022)	APTOS 2019 dataset	Diabetic Retinopathy	CNN CLAHE		The study applied Contrast Limited Adaptive Histogram Equalization (CLAHE) and Gaussian blurring techniques. Additionally, they subtracted the local mean color from images to enhance contrast, which improved the visibility of spots and blood vessels.	The preprocessing methods effectively enhanced image quality and improved feature visibility, contributing to the study's findings.	No proper results are tabulated
Syed Inthiyaz [10] (2023)	Xiangya-Derm[11]	Skin diseases	CNN	87%	The research involved resizing the images to extract the most relevant features from skin images.	Resizing was a crucial preprocessing step employed to highlight key features within the skin images, contributing to the study's success in achieving 87% accuracy.	Only six skin diseases are considered. The dataset is not standardized
Abdul Rafay [12] (2023)	Customized Dataset	Skin diseases	Efficient Net	87.15%	Augmentations	The study demonstrated that applying augmentations to the data significantly contributed to the improved accuracy of 87.15%.	
Maryam Naqvi [13] (2023)	Xiangya-Derm	Skin Cancer	Review paper		The research, presented in a review paper, emphasized the	The review paper highlighted the significance of various preprocessing	

					importance of augmentation, addressing imbalanced data, filtering, and segmentation techniques.	techniques for skin cancer analysis, including data augmentation, handling imbalanced data, filtering, and segmentation.	
Ghadah Alwakid [14] (2023)	HAM10000 Dataset[15]	Skin tumors	Finest Tuning with Inception-V3. Inception Resnet-V2.	89.7% 91.3%	Fine-tuning with Inception-V3 and InceptionResnet-V2 on the HAM10000 Dataset was performed.	The study showcased the importance of data augmentation techniques, which played a pivotal role in achieving an impressive accuracy of 89.7% and 91.3%.	
Chen, Min[16] (2020)	Customized dataset (6144 images)	Skin disease classification	LeNet-5 AlexNet VGG16	70% (Blackheads) 91% 87%	Standardization techniques were applied for skin disease classification using LeNet-5, AlexNet, and VGG16.	The study indicated that standardization was essential in obtaining reliable and consistent results, especially for classifying skin diseases.	
Pravin R. [17] Kshirsagar 2022	Customized dataset	Skin disease classification	MobileNet V2–LSTM	86.57%	To achieve high accuracy, the study proposed the fusion of segmentation with additional morphology marks.	The research suggested that applying morphology-based segmentation could enhance accuracy in skin disease classification.	
Keumsun Park[18] 2021	BIPED + Real time images[19]	Edge Detection	Canny Edge detection	80%	Canny Edge detection, Gaussian filtering, and data augmentation were employed, leading to a significant improvement in F1 results.	The study emphasized the impact of preprocessing techniques, particularly edge detection and data augmentation, which yielded substantially improved results.	
Rohith Kundu 2021 [20]	Kermany dataset RSNA[21]	Pneumonia detection in chest X-ray images	GoogLeNet, ResNet-18, and DenseNet-121	98.81% 5-fold cross-validation scheme	The study achieved an impressive accuracy of 98.81% using GoogLeNet, ResNet-18, and DenseNet-121 with a 5-fold cross-	The research highlights the potential for improving pneumonia detection in chest X-ray images through contrast enhancement	

			With	86.85%	validation scheme. Additionally, an ensemble scheme yielded an accuracy of 86.85%. To further enhance accuracy, contrast enhancement and lung image segmentation before classification are recommended.	and lung image segmentation as preprocessing steps.	
Gouda, [22] W (2022)	COV-PEN dataset [23] and CXR Images	Detection of COVID-19 Based on Chest X-rays	Resnet-50	99.63%	The study achieved a remarkable accuracy of 99.63% in detecting COVID-19 based on chest X-rays using Data augmentation and image enhancement technique to intensify CXR images and eliminate noise.	The research underscores the significance of data preprocessing techniques, including image enhancement and augmentation, in enhancing the accuracy of COVID-19 detection.	
Tanoy Debnath [24] (2022)	JAFFE and CK+ [25,26]	facial emotion recognition	LBP, ORB and CNN	92.05% 98.13%	In the context of facial emotion recognition using LBP, ORB, and CNN, normalization, gray scaling, and redimensioning were performed to ensure more reliable and efficient performance with the CNN approach.	The study emphasizes the importance of preprocessing techniques for achieving reliable and efficient facial emotion recognition with CNN models.	
A. Jaiswal[27] (2020)	Japanese Female Face Expression (JAFFE) FERC-2013	Facial Emotion Detection Using Deep Learning	CNN	70.14%	In the realm of facial emotion detection using CNN, the study achieved an average accuracy of 70.14%. Data transformation and normalization were	The research demonstrates the relevance of data preprocessing steps like data transformation and normalization in enhancing facial	

					integrated into the model.	emotion detection with CNN.	
Swadha Gupta [28]	FER-2013[29], CK+ and RAF-DB	Facial Emotion Detection Using Deep Learning	Deep Learning	78.8%.	89.11%, 90.14% and 92.32% for Inception-V3, VGG19 and ResNet-50	The proposed system was tested on 20 learners in an online learning scenario, and it correctly detected the "engaged" and "disengaged" states based on automatic facial emotion recognition. The proposed approach has also outperformed the existing work's methods	
Rodiah [2021] [30]	DRIVE[31]	Retina identification	Neural network	97.5%	For retina identification with neural networks, preprocessing was applied to convert retina images to grayscale values and extract input features. This resulted in a recognition accuracy of 98% for image patterns. 10 times of randomized trials and resulted in 9 correct identifications	Preprocessing, involving the segmentation. The resulted image is transformed using rotation, enlargement, shifting, cutting and reversing to increase the quantity of the samples	
M Usman Akram 1 [32] (2020).	Retina Identification Data Base (RIDB) and DRIDB	retina based person identification	Without using minutiae points Segmentation	92.50% Vascular 100 % and Non-vascular 92.5%	The proposed method utilized both vascular and non-vascular features for identification and yields recognition rates of 100 % and 92.5% respectively.	The study demonstrates the effectiveness of usage of camera TOPCON-TRC 50 EX plays a major role in extraction in non-vascular retina recognition.	

Lukáš Semerád [33]	Messidor [34], e-ophtha [35], High-Resolution Fundus (HRF) [36] and Retina EBD STRaDe (EBD).	To locate the individual bifurcations and crossings in the retinal image	Euclidian Distance	Same Eye marked by different person 86.50 Same Eye marked by Same person 93.50%	The main part was, of course, based on a comparison of the locations of the points in both images. Another part of the principle was based on a set of almost 1000 manually marked images where all bifurcations and crossings were located.	he evaluation algorithm was illustrative only to show how the individual parts worked.	
Awais Salman Qazi (2022) [37]	Mixed datasets Cross dataset CK+ and JAFFE	Emotion Detection Using Facial Expression	CNN	92.66% 94.94%.	Emotion detection using facial expression was carried out with CNN. The study achieved an accuracy of 92.66% and 94.94% through various preprocessing steps, including face detection, cropping, flipping, and angle sampling.	including face detection, cropping, flipping, and angle sampling	
Sunil S Harakannanavar [38] (2022)	JAFFE, Cohn – Kanade, Extended Cohn – Kanade, MMI, MUG, Taiwanese Facial Expression, Yale, AR face database	Emotion Detection Using Facial Expression	Fusion-HoG +LBP+ FKBD at feature level. SVM KNN	98.26% 96.51%	Emotion detection using facial expression was accomplished using a Fusion-HoG + LBP + FKBD approach, along with SVM and KNN classifiers. Cropping and scaling were applied as preprocessing steps, resulting in accuracies of 98.26% and 96.51%.	The study emphasizes the importance of preprocessing, including cropping and scaling, in enhancing emotion detection accuracy.	
Speech Processing							

Margaret Lech[39] (2020)	Berlin Emotional Speech (EMO-DB)[40]	Real-Time Speech Emotion Recognition	AlexNet	average accuracy of 82% (7 emotions) 79% 75%	The study achieved an average accuracy of 82% for recognizing seven emotions. Data preprocessing involved bandwidth reduction from 8kHz to 4kHz and companding	The study demonstrates that reducing the bandwidth and applying companding techniques can contribute to improved speech emotion recognition accuracy.	
Ala Saleh Alluhaidan [41] (2023)	Emo-DB, SAVEE[42], and RAVDESS datasets [43]	Speech Emotion Recognition	CNN	97%, 93%, and 92%	Utilizing multiple datasets, the study achieved high accuracy using a CNN model. Data preprocessing involved a high-pass filter called Finite Impulse Response (FIR).	The use of a high-pass filter like FIR in the preprocessing phase contributed to the study's high accuracy in speech emotion recognition.	
Aouani, H.,(2020)[44]	RML dataset	Speech Emotion Recognition	Support Vector Machines (SVM)	Kernel Linear 55.5 Kernel Polynomial 64.19 Kernel RBF 65.43	The study utilized the RML dataset and Support Vector Machines (SVM). A 42-D vector of audio features, including 39 MFCC, Zero Crossing Rate (ZCR), Harmonic to Noise Rate (HNR), and Teager Energy Operator (TEO), was used.	The combination of multiple audio features in the preprocessing phase contributed to varying SVM kernel performance, demonstrating the importance of feature selection in speech emotion recognition.	
Jagjeet Singh [45] (2023)	RAVDESS, SAVEE, and TESS datasets. 440, 1920, and 2800 speech files	Speech Emotion Recognition	CNN-2D + LSTM + Attention	57.50, 74.44, and 99.81% 90.19% (Combined Dataset –)	Multiple datasets were combined, and a CNN-2D + LSTM + Attention model achieved impressive accuracy. Data preprocessing involved normalization and augmentation.	The use of data normalization and augmentation techniques played a key role in achieving high accuracy, especially when working with combined datasets.	

Liu, G., Cai, S., & Wang, C. (2023).[46]	Interactive Emotional Dyadic Motion Capture (IEMOCAP) [47]	Speech Emotion Recognition	CNN	Average weighted81%	The study utilized the Interactive Emotional Dyadic Motion Capture (IEMOCAP) dataset and CNN. The preprocessing involved using MFCC.	The study achieved an average weighted accuracy of 81%, showcasing the effectiveness of MFCC-based preprocessing for emotion recognition.	
Beenaa Salian (2021)[48] ,	RAVDES Surrey Audio-Visual Expressed Emotion(SAVEE) Toronto emotional speech set (TESS)	Speech Emotion Recognition	Time Distributed CNN and LSTM	89.26%	Mel Spectrograms used for feature extraction	The audio samples are uniformaly splitted to remove noise and to reduce the size of the file	
Sourabh Suke(2021) [49]	RAVDES Toronto Emotional Speech Set (TESS) Datase	Speech Emotion Recognition System	Convolutional neural networks (CNNs), (three layers) 500 Epochs	85.71%	The study used the RAVDES and TESS datasets and employed CNNs with three layers for 500 epochs. Preprocessing techniques involved using Mel Frequency Cepstral Coefficient (MFCC).	The study achieved a recognition rate of 85.71%, demonstrating the effectiveness of MFCC-based preprocessing in speech emotion recognition.	
Hadhami(2020) [50]	RML database	Speech Emotion Recognition System	SVM	70.37%	The study utilized the RML database and employed SVM. Feature dimension reduction was applied in the preprocessing phase.	The study highlights the importance of feature dimension reduction as a preprocessing step for improving the accuracy of speech emotion recognition with SVM.	
D. Siddhart (2021) [51]	Customize dataset (few images)	Edge Detection Technique	ANN + Kalman Filtering	Canny is compared with Canny	This study focused on edge detection techniques using an ANN and Kalman Filtering. Preprocessing	The study demonstrates the significance of preprocessing with various filters,	

				,sobel,prewitt	involved comparing different edge detection methods, including Canny, Sobel, and Prewitt.	especially in edge detection tasks.	
Text Mining							
S. M. Sadjadi(2021) [52]	Reuters-21578 news dataset BBC News documentation, 2,225 news documents published from 2004 to 2005,	Clustering Text documents	Semi supervised Model And word2Vec	For 600 concepts 77.13%	The study used the Reuters-21578 news dataset and BBC News documentation for clustering text documents. Data preprocessing involved tokenization after removing stop-words and pre-processing.	The study achieved an accuracy of 77.13% in clustering text documents using a semi-supervised model and Word2Vec embeddings. The preprocessing step of removing stop-words and tokenization contributed to this result.	
van Vliet L (2020)[53]	Twitter Parliamentarian Database (TPD)]53]	Assessment of New Forms of politics across 26 countries	Page Rank	Identification of clusters by varying cluster parameters	The study assessed using $\chi2$ and Cramer's V measures, wherein Cramer's V show the strength of that relationship,	The study employed clustering algorithms to identify patterns in social media data related to political organization. Preprocessing is done automatically by identify and verifying twitter accouts	
Z. Dorrani [54] (2020)	Berkley Segmentation Dataset[55]	Image Edge Detection	Fuzzy Ant Colony Optimization Algorithm	fuzzing, we are able to improve the behavior of the algorithm	The study focused on image edge detection using the Berkeley Segmentation Dataset. Data preprocessing utilized the Fuzzy Ant Colony Optimization Algorithm and standardization.	The study improved the behavior of the edge detection algorithm through the preprocessing technique of standardization. Fuzzing was also employed to enhance algorithm performance.	

Pham, B [55] (2021)	Articles	Abstract Screening	SVD LDA	88% /89% sensitivity, 99% /99% specificity, 71% /72% precision, and F1-score of 79% /79%, 98% /97% accuracy	The study conducted abstract screening of articles using SVD and LDA techniques. Data preprocessing included tokenization, lemmatization, parts-of-speech tagging, and semantic annotation as needed.	The study achieved high sensitivity, specificity, precision, and accuracy in abstract screening. The preprocessing steps, including tokenization and semantic annotation, likely contributed to these results.	
Arafat Hossain[56](2021)	Bangladesh's daily newspaper headlines, The Daily Star	Sentiment Analysis of Newspaper Headlines	Normal distance and Ward method of clustering	Identified most repeated words in the news paper	The study performed sentiment analysis of Bangladesh's daily newspaper headlines from The Daily Star. Data preprocessing included using the normal distance and Ward method of clustering and creating a corpus of the data by removing punctuation, numbers, and other unnecessary elements.	The study successfully identified sentiment in newspaper headlines. Preprocessing, which involved cleaning the text data by removing punctuation and numbers, was essential for meaningful analysis.	
Genetic Algorithm Based Machine Learning Technique							

Mr. Sanghyeop Lee [57] (2021)	PET/CT [58]	Alzheimer's disease detection	Convolutional Neural Networks Genetic Algorithm (GA)	81.74 which is greater than 70.01 with genetic CNN(11.73%)	The study focused on Alzheimer's disease detection using PET/CT scans. Data preprocessing included the use of Convolutional Neural Networks (CNN), Genetic Algorithm (GA), and various image processing techniques in MATLAB, such as normalization, resizing from 3D to 2D, and using the SPM12 toolbox.	The combination of CNN and GA, along with the preprocessing steps, led to an accuracy of 81.74% in Alzheimer's disease detection, surpassing previous results.	
Delia Dumitru [59] (2021)	Brodatz dataset MRI dataset	Edge Detectors for Medical Images	PSO optimization and transfer learning	proposed method performed better than Canny on average on our cardiac	The study involved the use of edge detectors for medical images, specifically the Brodatz dataset and MRI data. Preprocessing included PSO optimization, transfer learning, and the application of Gaussian filters.	The proposed method outperformed the Canny edge detector on average for cardiac images, demonstrating the effectiveness of the preprocessing techniques.	
[60] Anu Saini	Hotel review dataset	Speech emotion recognition	Genetic algorithm to DTW	multinominal Naive Bayes (MNB), logistic regression (LR), and LSVM	n-gram and term frequency-inverse document frequency (TFIDF) approach to extract features in computation. MNB has an average accuracy of eighty seven percent (87%), LR has a ninty five percent (95%) average accuracy, and LSV has a ninety six percent	linear support vector machine (LSVM played a crucial role in optimizing systems for speech recognition, highlighting their potential in preprocessing for speech emotion recognition.	

					(96%) average accuracy		
Gurpreet [61] Kaur(2018)	Customized dataset with thousands of words	Speech Recognition	DNN and genetic MFCC+GA +DNN	97.19%(one iteration	The study employed a customized dataset with thousands of words for speech recognition. Preprocessing involved deep neural networks (DNN), genetic algorithms for feature extraction (MFCC+GA), and noise reduction.	The results demonstrated that MFCC, optimized with genetic algorithms, offered the best performance in clean and noisy environments.	
Fuzzy Based Machine Learning Techniques							
M. R. [62] Dileep(2019)	Customized dataset 1000 700 training 300 Testing	Human FacialExpression	Fuzzy Inference system	95%	The study utilized a customized dataset consisting of 1000 samples for training and 300 for testing for human facial expression recognition. Data preprocessing included normalizing the images to a size of 64 × 64 pixels.	The study achieved an impressive 95% accuracy in emotion recognition using a fuzzy inference system. The preprocessing step of resizing the images to a uniform size likely contributed to the model's success in recognizing facial expressions.	
Facial expression identification using Machine Learning Techniques							
TS, Ashwin & Guddeti, Rammohana Reddy (2020)	The authors created a dataset of over 8000 single face in a single image frame and 12000 multiple faces in a single image frame	Automatic detection of students' affective states in classroom environment	Hybrid convolutional neural networks	86% for posed images and 70% spontaneous affective states	The key to the robust deep learning model is the high quality data. But, it is a challenge to obtain such data. Hence augmentation of datasets is done to address the issue.	The study predicts the students' emotional and behavioral engagement separately in both e-learning and classroom environments.	

Sarra Ayouni , Fahima Hajjej, Mohamed Maddeh, Shaha Al-Otaibi	Data of 360 students enrolled in College of Computer and Information Sciences at PNU.	A new ML-based approach to enhance student engagement in online environment	Decision Tree, Support Vector Machine and Artificial Neural Network	Artificial Neural Network (85%), Support Vector Machine (80%) and Decision Tree (75%)	This study shows that measuring students' level of engagement and providing them with feedback at the appropriate moment ensure students focus on the course, which will in turn support student performance and experience.	Because of the complexity of student engagement construct, many other indicators such as course design, teaching style and other factors external to the course should be investigated.	
S. Zhang, X. Pan, Y. Cui, X. Zhao and L. Liu,	BAUM-1s, RML database, MMI database	Learning Affective Video Features for Facial Expression Recognition	FER is used in video sequences via a hybrid deep learning model	55.85% on the BAUM-1s dataset, 73.73% on the RML dataset, and 71.43% on the MMI dataset	A spatial CNN network processing static frame-level cropped facial images and a temporal CNN network processing optical flow images produced between consecutive frames.	VGG16 model is pre-trained on ImageNet data to individually fine-tune the spatial CNN network and the temporal CNN network on target video-based facial expression data. To deeply fuse the learned spatio-temporal CNN features, a deep DBN model is trained to jointly learn discriminative spatio-temporal features.	

Table 1: Outcomes of data preprocessing techniques applied to a range of datasets.

IMAGE PRE-PROCESSING TECHNIQUES AND THE MULTIFACETED WORLD OF DIGITAL IMAGERY

Images in the digital world are vital for our visual perception. They are created by converting real-world visuals into digital form, where each pixel encodes visual information. These digital images are essentially two-dimensional arrays of pixels, with each pixel representing specific information about its position and intensity. Digital images are typically created by scanning documents through devices like scanners, which measure reflected light and convert the measurements into binary digits. There are two primary categories of digital imagery: bitmaps and vector images. Bitmaps are pixel-based and include formats like BMP, PNG, JPG, and GIF, while vector images use mathematical equations to represent lines and curves, making them infinitely scalable.

There are different types of digital images based on their pixel attributes, such as binary images with only black and white values, grayscale images with various shades of gray, and color images that use red, green, and blue (RGB) channels to create rich, detailed visual representations.

Indexed images use a colormap matrix for more control over color rendering.

Image analysis involves extracting valuable information from digital images, from counting objects to identifying shapes. The connection between image pre-processing techniques and image analysis ensures the quality and integrity of the extracted information. As technology advances, digital imagery continues to evolve, finding applications in various fields like medical imaging and art.

IMAGE ENHANCEMENT: UNVEILING HIDDEN DETAILS

Image enhancement is a crucial aspect of image processing, focusing on revealing hidden details and improving image quality. It involves adjustments to attributes like brightness and contrast to highlight specific features while reducing or removing irrelevant elements. This process can include noise reduction, detail sharpening, and contrast enhancement, depending on the task's requirements. Augmentation, a key component of data preprocessing, plays a vital role in enhancing the diversity of datasets to improve the robustness of machine learning models. Various transformation techniques are applied to expand the dataset, introducing variations and increasing its size for training. These transformations encompass flipping, shearing, zooming, and noise introduction. This diversifies the dataset, equipping the model to handle real-world objects under different conditions and preventing overfitting.

Flipping: Creates new image instances by horizontally or vertically altering the orientation, providing different perspectives for the model to learn from.

Shearing: Tilts images at various angles to introduce a range of viewpoints and perspectives, enriching the dataset.

Zooming: Expands the dataset by including images of different sizes, allowing the model to adapt to objects of varying scales.

Rescaling: Standardizes pixel values to improve model performance, typically by scaling them to the range [0,1].

Shear Range: Alters pixel directions to generate diverse transformed images, further enhancing dataset diversity.

Horizontal Flip: Generates horizontally mirrored images, broadening the model's ability to handle various orientations.

Addressing unbalanced data distribution within a dataset is crucial, as it can lead to biased results and poor model performance. The Synthetic Minority Over-sampling Technique (SMOTE) is a widely used oversampling approach that combats class imbalance by generating synthetic instances. SMOTE interpolates between instances of the minority class and their nearest neighbors, introducing diversity into the synthetic dataset and improving the model's generalization capabilities. However, it may not be suitable for cases where minority class instances are densely clustered or contain outliers.

RESULTS

Figure 1 presents the outcomes following the application of diverse preprocessing methods to the datasets specified in Table 2. The combined information from Table 2 and Figure 1 undeniably underscores the significance of preprocessing in enhancing the quality and relevance of data, whether it pertains to image or audio data of varying types.

6. FUTURE PROSPECTS

Future prospects in image processing and machine learning include:

Emerging Augmentation Strategies: Researchers have promising opportunities to explore advanced data augmentation techniques tailored to diverse datasets. Developing novel transformations, fusion strategies, and innovative augmentation methods based on learning principles can enrich the pool of lifelike and varied image variations, enhancing the learning process.

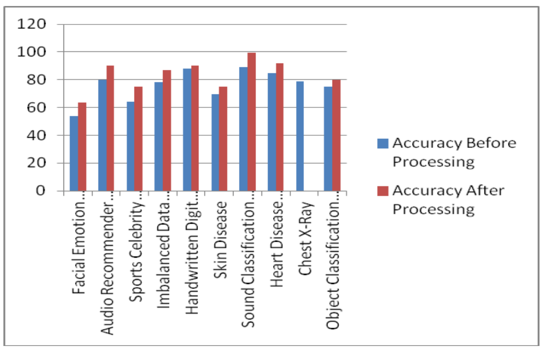


Fig 1: Application of diverse preprocessing methods

Crossing Domains: Investigating domain adaptation techniques is a valuable direction for future research. Especially when dealing with target domains that differ from the original dataset, these methods can significantly improve model performance. Transferring knowledge effectively from alternative datasets or domains to enhance models like CIFAR-10 is a promising avenue for exploration.

Navigating Noise Challenges: Dealing with noisy or imperfect images will be a significant focus in the future. Convolutional Neural Networks (CNNs) trained on clean datasets like CIFAR-10 may encounter difficulties when faced with real-world images marked by noise, occlusions, or distortions. Research should aim to develop techniques that enhance noise resilience and enable models to perform effectively in complex scenarios.

Adaptive Data Shaping: Future research can explore adaptive pre-processing techniques. Developing methodologies that dynamically adjust data transformations based on the unique attributes of input images can streamline data preparation and customization. These adaptive strategies can efficiently orchestrate various pre-processing steps to meet the specific needs of different image characteristics.

7. DISCUSSION

In this paper, we embarked on a comprehensive exploration of various pre-processing techniques, unraveling their roles and impacts across a spectrum of applications. The journey spanned from basic enhancements to more sophisticated interventions, all aimed at enhancing the quality, accuracy, and robustness of

TABLE 2 : VARIOUS DATASETS

Purpose	Dataset	Methodology	Accuracy Before Processing	Accuracy After Processing	Remarks
Facial Emotion Detection	FER2013	CNN	VA 53.72%	VA 63.72%	Improved accuracy through data normalization.
Audio Recommender System	GTZAN dataset	Cross Gradient Booster	80%	90%	Enhanced accuracy after data normalization.
Sports Celebrity Classification	Sports-Person-Classifier	SVM-Linear Kernel	64%	74.86%	Improved accuracy by cropping the facial region of the image.
Imbalanced Data Preprocessing	Thyroid sick	Logistic Regression	Precision 0.78 Recall 0.55	Precision 0.35 Recall 0.87 Precision 0.36 Recall 0.87	Addressed class imbalance through random oversampling and SMOTE.
Handwritten Digit Recognition	MNIST	KNN	Accuracy 88%	Accuracy 90	Enhanced accuracy through data cleaning, transformation, reduction, and feature selection.
Skin Disease	Dog breed dataset	ResNet-18	69.69%	75%	Improved accuracy via image augmentation, resizing, normalization, Gaussian blur, grayscale conversion, median blur, bilateral filtering, and contrast enhancement.
Sound Classification using ANN	Us8k	ResNet-152, ESC-10, DenseNet-161	89%, 85%, 80%	99.04%, 99.49%, 97.57%	Enhanced accuracy achieved by applying MFCC preprocessing and noise reduction.
Heart Disease Prediction	Heart-disease prediction dataset	Decision Tree	85%	92%	Improved accuracy after data cleaning, organization, and handling missing data.
Chest X-Ray	Pneumonia detection	CNN	79%	89. 92%	Increased accuracy through data augmentation, histogram equalization, Gaussian blur, and adaptive masking.
3D-Medical Imaging Preprocessing	Abdomen CT scans	CNN	Significant enhancements were realized in image quality, quantitative analysis, and the detection of subtle anomalies.		Improved image data quality through standardization, noise reduction, spatial alignment, intensity alignment, histogram equalization, cropping, padding, and resampling.
Object Classification using CNN	CIFAR-10 [85]	(CNNs)	75%	80%	Image Augmentation: Resizing: Normalization: Grayscale Conversion:

image-based models. Through meticulous analysis and experimentation, we demonstrated that pre-processing acts as a crucial cornerstone, shaping the foundations for superior performance.

Looking ahead, the future horizons beckon with prospects for advanced augmentation, seamless domain adaptation, noise mitigation, and adaptive pre-processing. These horizons promise to unravel new dimensions for the research community.

REFERENCES

1. Syed Inthiyaz, Baraa Riyadh Altahan, Sk Hasane Ahammad, V Rajesh, Ruth Ramya Kalangi, Lassaad K. Smirani, Md. Amzad Hossain, Ahmed Nabih Zaki Rashed,Skin disease detection using deep learning,Advances in Engineering Software,Volume 175, 2023, 103361, ISSN 0965-9978, https://doi.org/10.1016/j.advengsoft.2022.103361.

2. Khalid K. Al-jabery et al, in Computational Learning Approaches to Data Analytics in Biomedical Applications, 2020 Hardback ISBN: 9780128144824 eBook ISBN: 9780128144831

3. Desai, B., Paliwal, M., & Nagwanshi, K. K. Study on Image Filtering Techniques, Algorithm and Applications. (2022). ArXiv. /abs/2207.06481

4. Rafay, A., & Hussain, W. Efficient SkinDis: An Efficient Net-based classification model for a large manually curated dataset of 31 skin diseases. Biomedical Signal Processing and Control, 2023. 85, 104869. https://doi.org/10.1016/j.bspc.2023.104869

5. Tachibana Y, Obata T, Kershaw J, Sakaki H, Urushihata T, Omatsu T, Kishimoto R, Higashi T. The Utility of Applying Various Image Preprocessing Strategies to Reduce the Ambiguity in Deep Learning-based Clinical Image Diagnosis. Magn Reson Med Sci. 2020 May 1;19(2):92-98. doi: 10.2463/mrms.mp.2019-0021. Epub 2019 May 10. PMID: 31080211; PMCID: PMC7232029.

6. Maharana, K., Mondal, S., & Nemade, B. A review: Data pre-processing and data augmentation techniques.Global Transitions Proceedings, 2022.3(1), 91-99. https:/doi.org/10.1016/j.gltp.2022.04.020

7. Swagata Boruah Computers, Materials & Continua 2023, 75(1), 927-942. doi.org/10.32604/cmc.2023.035143

8. APTOS 2019 Blindness Detection, 2019. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection/. [Accessed: 10-Jun-2019]

9. Bakhshi A, Hajizadeh K, Tanhayi MR, et al. Diabetic retinopathy diagnosis using image processing methods. Adv Obes Weight Manag Control 2022,12(5):132-134. DOI: 10.15406/aowmc.2022.12.00375

10. Syed Inthiyaz, Baraa Riyadh Altahan, Sk Hasane Ahammad, V Rajesh, Ruth Ramya Kalangi, Lassaad K. Smirani, Md. Amzad Hossain, Ahmed Nabih Zaki Rashed, Skin disease detection using deep learning, Advances in Engineering Software, Volume 175,

2023, 103361, ISSN 0965-9978, https://doi.org/10.1016/j.advengsoft.2022.103361.(https://www.sciencedirect.com/science/article/pii/S0965997822002629)

11. Bin Xie , XiangyaDerm: A Clinical Image Dataset of Asian Race for Skin Disease Aided Diagnosis, https://airl.csu.edu.cn/PDFs/LABELS2019_XiangyaDerm.pdf

12. Abdul Rafay, Waqar Hussain EfficientSkinDis: An EfficientNet-based classification model for a large manually curated dataset of 31 skin diseases. Biomed. Signal Process. Control. 85: 104869 (2023)

13. Naqvi, M.; Gilani, S.Q.; Syed, T.; Marques, O.; Kim, H.-C. Skin Cancer Detection Using Deep Learning—A Review. Diagnostics 2023, 13, 1911. https://doi.org/10.3390/diagnostics13111911

14. Alwakid G, Gouda W, Humayun M, Jhanjhi NZ. Diagnosing Melanomas in Dermoscopy Images Using Deep Learning. Diagnostics (Basel). 2023 May 22;13(10):1815. doi: 10.3390/diagnostics13101815. PMID: 37238299; PMCID: PMC10217211.

15. Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci Data 5, 180161 (2018). https://doi.org/10.1038/sdata.2018.161

16. Min Chen et.al , AI-Skin : Skin Disease Recognition based on Self-learning and Wide Data Collection througha Closed Loop Framework,arXiv:1906.01895[cs.CV] https://doi.org/10.48550/arXiv.1906.01895

17. Kshirsagar, P.R.; Manoharan, H.; Shitharth, S.; Alshareef, A.M.; Albishry, N.; Balachandran, P.K. Deep Learning Approaches for Prognosis of Automated Skin Disease. Life 2022, 12, 426. https://doi.org/10.3390/life12030426

18. Park K, Chae M, Cho JH. Image Pre-Processing Method of Machine Learning for Edge Detection with Image Signal Processor Enhancement. Micromachines (Basel). 2021 Jan 11;12(1):73. doi: 10.3390/mi12010073. PMID: 33440903; PMCID: PMC7827319.

19. https://www.kaggle.com/datasets/xavysp/biped

20. Kundu R, Das R, Geem ZW, Han G-T, Sarkar R (2021) Pneumonia detection in chest X-ray images using an ensemble of deep learning models. PLoS ONE 16(9): e0256630. https://doi.org/10.1371/journal.pone.0256630

21. Kermany D., Zhang K. & Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. (Mendeley,2018)

22. Alwakid, G et.al, Diagnosing Melanomas in Dermoscopy Images Using Deep Learning. Diagnostics, 13(10), 1815. https://doi.org/10.3390/diagnostics13101815

23. Cohen J.P., Morrison P., Dao L., Roth K., Duong T.Q., Ghassemi M. Covid-19 image data collection: Prospective

predictions are the future. arXiv. 20202006.11988

24. Debnath, T., Reza, M. M., Rahman, A., Beheshti, A., & Band, S. S. (2022). Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity. Scientific Reports, 12(1), 1-18. https://doi.org/10.1038/s41598-022-11173-0

25. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

26. Lyons, Michael, Kamachi, Miyuki, & Gyoba, Jiro. (1998). The Japanese Female Facial Expression (JAFFE) Dataset [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3451524

A. Jaiswal, A. Krishnama Raju and S. Deb, "Facial Emotion Detection Using Deep Learning," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9154121.

27. Gupta, S., Kumar, P. & Tekchandani, R.K. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. Multimed Tools Appl 82, 11365–11394 (2023). https://doi.org/10.1007/s11042-022-13558-9

28. L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim and S. Bahri Musa, "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi," 2020 Fifth International Conference on Informatics and Computing (ICIC), Gorontalo, Indonesia, 2020, pp. 1-9, doi: 10.1109/ICIC50835.2020.9288560.

29. Rodiah, Madenda S, Susetianingtias DT, Fitrianingsih, Adlina D, Arianty R. Retinal biometric identification using convolutional neural network. Computer Optics 2021; 45(6): 865-872. DOI: 10.18287/2412-6179-CO-890.

30. https://www.kaggle.com/datasets/andrewmvd/drive-digital-retinal-images-for-vessel-extraction

31. Akram, M. & abdul salam, Anum & Khawaja, Sajid & Naqvi, Syed & Khan, Shoab. (2020). RIDB: A Dataset of fundus images for retina based person identification. Data in Brief. 33. 106433. 10.1016/j.dib.2020.106433

32. Lukáš Semerád and Martin Drahanský Retina Recognition Using Crossings and Bifurcations Applications of Pattern Recognition Edited by Carlos M. Travieso-Gonzalez DOI: 10.5772/intechopen.96142

33. Qazi, A. S., Farooq, M. S., Rustam, F., Villar, M. G., Rodríguez, C. L., & Ashraf, I. Emotion Detection Using Facial Expression Involving Occlusions and Tilt. Applied Sciences, 12(22), 11797. https://doi.org/10.3390/app122211797

34. Decencière E, et al. TeleOphta: Machine learning and image processing

methods for teleophthalmology. IRBM (2013), http://dx.doi.org/10.1016/j.irbm.2013.01.010

35. High-Resolution Fundus (HRF) Image Database. https://www5.cs.fau.de/research/data/fundus-images/. Accessed August 27, 2020

36. Qazi, A.S.; Farooq, M.S.; Rustam, F.; Villar, M.G.; Rodríguez, C.L.; Ashraf, I. Emotion Detection Using Facial Expression Involving Occlusions and Tilt. Appl. Sci. 2022, 12, 11797. https://doi.org/10.3390/app122211797

37. Harakannanavar SS, Sapnakumari C, Ramachandra AC, Pramodhini R, Prashanth CR (2023) Performance Evaluation of Fusion Based Efficient Algorithm for Facial Expression Recognition. Indian Journal of Science and Technology 16(4): 266-276. https://doi.org/ 10.17485/IJST/v16i4.1891

38. Margaret Lech et.al, Real-Time Speech Emotion Recognition Using a Pre-trained Image Classification Network: Effects of Bandwidth Reduction and Companding ,Front. Comput. Sci., 26 May 2020 Sec. Human-Media Interaction Volume 2 - 2020 https://doi.org/10.3389/fcomp.2020.00014

39. https://www.tu.berlin/en/kw/research/projects/emotional-speech

40. Alluhaidan, A.S.; Saidani, O.; Jahangir, R.; Nauman, M.A.; Neffati, O.S. Speech Emotion Recognition through Hybrid Features and Convolutional Neural Network. Appl. Sci. 2023, 13, 4750. https://doi.org/10.3390/app13084750

41. http://kahlan.eps.surrey.ac.uk/savee/

42. https://www.kaggle.com/datasets/ryersonmultimedialab/ryerson-emotion-database

43. Aouani, H., Ayed, Y.B., 2020. Speech emotion recognition with deep learning. Procedia Computer Science 176, 251–260.

44. Singh, J.; Saheer, L.B.; Faust, O. Speech Emotion Recognition Using Attention Model. Int. J. Environ. Res. Public Health 2023, 20, 5140. https://doi.org/10.3390/ijerph20065140

45. Liu, G., Cai, S. & Wang, C. Speech emotion recognition based on emotion perception. J AUDIO SPEECH MUSIC PROC. 2023, 22 (2023). https://doi.org/10.1186/s13636-023-00289-4

46. C. Busso, M. Bulut, C.C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J.N. Chang, S. Lee, and S.S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," Journal of Language Resources and Evaluation, vol. 42, no. 4, pp. 335-359, December 2008.

47. Beenaa Salian et al., Speech Emotion Recognition using Time Distributed CNN and LSTM, ITM Web of Conferences 40, 03006 (2021), ICACC-2021, https://doi.org/10.1051/itmconf/20214003006

48. Siddiqui, M.F.H.; Dhakal, P.; Yang, X.; Javaid, A.Y. A Survey on Databases for Multimodal Emotion Recognition and an Introduction to the VIRI (Visible and InfraRed Image) Database. Multimodal

Technol. Interact. 2022, 6, 47. https://doi.org/10.3390/mti6060047

49. Hadhami Speech Emotion Recognition with deep learning, January 2020 Procedia Computer Science 176(4):251-260 DOI: 10.1016/j.procs.2020.08.027

50. D. Siddhartha , D. K. J. Saini*b, P. Singh, An Efficient Approach for Edge Detection Technique using Kalman Filter with Artificial Neural Network, International Journal of Engineering, Vol. 34, No. 12, (December 2021) 2604-2610

51. S. M. Sadjadi, H. Mashayekhi*, H. Hassanpour International Journal of Engineering IJE TRANSACTIONS C: Aspects Vol. 34, No. 12, (December 2021) 2648-2657

52. van Vliet L, Törnberg P, Uitermark J (2020) The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries. PLoS ONE 15(9): e0237073. https://doi.org/10.1371/journal.pone.0237073

53. Z. Dorrani Image Edge Detection with Fuzzy Ant Colony Optimization Algorithm Volume 33, Issue 12, December 2020, Pages 2464-2470 Z.

54. Pham, B., Jovanovic, J., Bagheri, E. et al. Text mining to support abstract screening for knowledge syntheses: a semi-automated workflow. Syst Rev 10, 156 (2021). https://doi.org/10.1186/s13643-021-01700-x

55. Hossain, A., Karimuzzaman, M., Hossain, M. M., & Rahman, A. (2021). Text Mining and Sentiment Analysis of Newspaper Headlines. Information, 12(10), 414. https://doi.org/10.3390/info12100414

56. Lee S, Jung JH, Kim D, Lim HK, Park MA, Kim G, So M, Yoo SK, Ye BS, Choi Y, Yun M. PET/CT for Brain Amyloid: A Feasibility Study for Scan Time Reduction by Deep Learning. Clin Nucl Med. 2021 Mar 1;46(3):e133-e140. doi: 10.1097/RLU.0000000000003471. PMID: 33512838.

57. Gatidis S, Hepp T, Früh M, La Fougère C, Nikolaou K, Pfannenberg C, Schölkopf B, Küstner T, Cyran C, Rubin D. A whole-body FDG-PET/CT Dataset with manually annotated Tumor Lesions. Sci Data. 2022 Oct 4;9(1):601. doi: 10.1038/s41597-022-01718-3. PMID: 36195599; PMCID: PMC9532417

58. Dumitru, D.; Dioșan, L.; Andreica, A.; Bálint, Z. A Transfer Learning Approach on the Optimization of Edge Detectors for Medical Images Using Particle Swarm Optimization. Entropy 2021, 23, 414. https://doi.org/10.3390/e23040414

59. Anu Saini , An investigation of machine learning techniques in speech emotion recognitionIndonesian Journal of Electrical Engineering and Computer Science 29(2):875 ,DOI:10.11591/ijeecs.v29.i2.pp875-882

60. Gurpreet Kaur1, 2, Mohit Srivastava3, and Amod Kumar https://doi.org/10.26636/jtit.2018.119617 Journal of Telecommunications and information technology

61. M. R. Dileep(2019), The Quantification of Human Facial Expression Using Trapezoidal Fuzzy Membership

Function, Recent Trends in Image Processing and Pattern Recognition: Second International Conference, RTIP2R 2018, Solapur, India, December 21–22, 2018, Revised Selected Papers, Part II 2

62. Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data. Frontiers in Energy Research, 9, 652801. https://doi.org/10.3389/fenrg.2021.652801

63. Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 3(1), 91-99. https://doi.org/10.1016/j.gltp.2022.04.020

64. Abbaschian, B.J.; Sierra-Sosa, D.; Elmaghraby, A. Deep Learning Techniques for Speech Emotion Recognition, from Databases to Models. Sensors 2021, 21, 1249. https://doi.org/10.3390/s21041249

65. Pulatov I, Oteniyazov R, Makhmudov F, Cho YI. Enhancing Speech Emotion Recognition Using Dual Feature Extraction Encoders. Sensors (Basel). 2023 Jul 24;23(14):6640. doi: 10.3390/s23146640. PMID: 37514933; PMCID: PMC10383041.

66. Hossain, A., Karimuzzaman, M., Hossain, M. M., & Rahman, A. (2021). Text Mining and Sentiment Analysis of Newspaper Headlines. Information, 12(10), 414. https://doi.org/10.3390/info12100414

67. Kotyrba, M.; Volna, E.; Habiballa, H.; Czyz, J. The Influence of Genetic Algorithms on Learning Possibilities of Artificial Neural Networks. Computers 2022, 11, 70. https://doi.org/10.3390/computers11050070

68. Lyimo, N.N.; Shao, Z.; Ally, A.M.; Twumasi, N.Y.D.; Altan, O.; Sanga, C.A. A Fuzzy Logic-Based Approach for Modelling Uncertainty in Open Geospatial Data on Landfill Suitability Analysis. ISPRS Int. J. Geo-Inf. 2020, 9, 737. https://doi.org/10.3390/ijgi9120737

69. Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M. Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez, Francisco Herrera, Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence, Information Fusion, Volume 99, 2023, 101805, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2023.101805.

70. Vaishali Kalra et al Importance of Text Data Preprocessing & Implementation in RapidMiner, The First International Conference on Information Technology and Knowledge Management January 2018 DOI: 10.15439/2017KM46

71. Amoako-Yirenkyi, P. , Appati, J. and Dontwi, I. (2016) Performance Analysis of Image Smoothing Techniques on a New Fractional Convolution Mask for Image Edge Detection. Open Journal of Applied Sciences, 6, 478-488. doi: 10.4236/ojapps.2016.67048.

72. Nitesh V. Chawla SMOTE: Synthetic Minority Over-sampling Technique Journal of Artificial Intelligence Research 16 (2002) 321–357

73. T. S., A., Guddeti, R.M.R. Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. Educ Inf Technol 25, 1387–1415 (2020). https://doi.org/10.1007/s10639-019-10004-6

74. Sarra Ayouni, Fahima Hajjej, Mohamed Maddeh, Shaha Al-Otaibi. A new ML-based approach to enhance student engagement in online environment. Plos One, 2021, https://doi.org/10.1371/journal.pone.0258788

75. S. Zhang, X. Pan, Y. Cui, X. Zhao and L. Liu, "Learning Affective Video Features for Facial Expression Recognition via Hybrid Deep Learning," in IEEE Access, vol. 7, pp. 32297-32304, 2019, doi: 10.1109/ACCESS.2019.2901521.