IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AUDIO-VISUAL E-WASTE CLASSIFICATION WITH A TWO-STREAM DEIT MODEL

[1]Asmi Jain, [2]Kotla Aishwarya, [3]Divya Reddy M, [4]Mona

[1]Student, [2]Student, [3]Student, [4]Assistant Professor
[1]Computer science and engineering,
[1]BNM Institute of technology, Bengaluru, Karnataka, India

Abstract: Electronic garbage, known as "e-waste," poses a significant environmental and public health risk. The World Bank estimates that highly industrial waste emissions total around 1.5 billion tons per year. Effective classification strategies are crucial due to its complicated and often hazardous nature. This study explores effective e-waste classification methods that prioritize sustainability, cost-effectiveness, and innovation. The suggested methods range from NVIDIA Jetson Nano-based deep learning models for identifying precious metals to thermal imaging techniques for classifying recyclable materials. Other methods include Faster RCNN, YOLO, and CNN with different architectures. The project also investigates the potential of computer vision technologies for e-waste.

***Index Terms* - NVIDIA's Jetson Nano, thermal imaging, Faster RCNN, YOLO, CNN, DEMATEL, ISM**

I. INTRODUCTION

The increasing amount of electronic waste, or "e-waste," is a serious threat to the environment and a public health issue. India, has emerged as the world's second largest mobile market, is also the fifth largest producer of e-waste, discarding roughly 18.5 lakh tons of electronic waste each year [13][14]. According to estimates from the World Bank, the annual global emissions of highly industrial waste are forecast to be 1.5 billion tons; by 2025, that amount is predicted to increase to 2.2 thousand pounds [8]. According to the United Nations Environment Programme (UNEP), E-waste is one of the fastest growing wastes globally with an annual growth of 3-4% in the world widely and it is estimated annual accumulation of around 74 million tons by the year 2030 [9]. Robust and effective categorization systems are essential for safe disposal, resource recovery, and environmental protection due to the complex and frequently dangerous nature of the material. detailed information about the classification of e-waste, including an examination of the technology and procedures used today and the difficulties in achieving precise and effective sorting. There are several difficulties in classifying and separating different e-waste components because of the wide variations in size, shape, texture, and material composition of electronic devices, advanced recognition systems that can distinguish subtle differences between things are required.

E-waste poses severe risks to human health, ecosystems, and the environment due to toxic compounds. Inadequate waste management contributes to pollution, affecting groundwater, waterways, and oceans, while informal disposal leads to air pollution. The health impacts on workers, including cardiovascular issues, birth defects, and cancer, are alarming, with women and children constituting a significant part of the informal sector. Despite increasing research on health effects, challenges in data collection persist, including diverse outcomes and the long latency period of diseases associated with E-waste exposure [10].

Conventional sorting techniques, which depend on human hands and eyes, frequently fail to handle the complex makeup of e-waste. In order to ensure proper recycling, disposal, and recovery of valuable components while avoiding environmental pollution and health concerns associated with mishandling these electronic devices, proper classification of e-waste is essential. Acknowledging these constraints, scientists have set out on an exciting search for cutting-edge technology that can transform the way e-waste is classified and transform this environmental burden into a useful resource.

Using thermal imaging techniques to classify e-waste is one promising approach. Based on their thermal signatures, researchers have created methods for categorizing different e-waste materials, including plastics, metals, glass, and printed circuit boards (PCBs).

With the help of this non-intrusive categorization method, items can be identified without coming into touch with one another. However, hardware prices may be an obstacle to the widespread use of thermal imaging, requiring more research into affordable alternatives. Its promise is demonstrated by research conducted at the Indian Institute of Technology Bombay, where it was able to separate e-waste components with an astounding 80% accuracy.

Deep learning models are becoming more and more potent instruments for classifying e-waste in parallel. Using NVIDIA's Jetson Nano development kit to create a deep learning model is one noteworthy example. The objective of this approach is to categorize waste materials into groups according to whether precious or non-precious metals are present. Even though these models might be accurate, there are issues with the inconvenience of doing processes remotely and the accompanying costs of heavy equipment and gear. The model's accuracy is higher than 92%.

Using computer vision technology (CVT) to improve the planning of e-waste collection. This method makes use of convolutional neural networks (CNNs) and region-based CNNs (R-CNNs) to enable the detection and classification of e-waste using photography. Though encouraging, it has drawbacks in that it only classifies a small number of waste categories and assumes that equipment is not seriously damaged [5].

To evaluate the factors that support CVT for sustainable e-waste management, decision-making instruments like Interpretive Structural Modelling (ISM) and the Decision-Making Trial and Evaluation Laboratory (DEMATEL) are used. These approaches shed light on the technological, organizational, legal, and infrastructure difficulties related to the early stages of CVT implementation in e-waste management.

Contrastive learning is tested as a part of the Faster RCNN for printed circuit board (PCB) electronic component detection. The objective of this investigation is to improve the precision of identifying electronic components in e-waste materials. Nevertheless, obstacles continue to exist, such as the requirement for large datasets and annotated labels in order to surpass conventional models.

Finally, Convolutional Neural Networks (CNNs) are used in the use of transfer learning in waste categorization to classify datasets into discrete waste types. Although this method increases accuracy, it also shows that more deployment optimizations are needed, possibly bringing the application to mobile platforms.

To sum up, these many approaches highlight the ever-changing field of e-waste classification research and the interdisciplinary character of efforts to address this urgent environmental concern. Every technique, from computer vision and deep learning to thermal imaging and decision-making tools, adds its own insights and difficulties to the search for efficient and sustainable e-waste treatment techniques. This study aims to investigate the possible synergies between developing technologies and propose strategies to integrate them into a comprehensive and flexible system that can address the always changing complications of e-waste, as technology continues to advance.

The subsequent sections are arranged in the following manner. Section II discusses the related works. Section III discusses the proposed system. Section IV demonstrates the results and discussion. Finally, Section V concludes the paper.

II. Related Works

An unprecedented period of connectedness and invention has been brought about by the rapidly expanding proliferation of electronic devices. But with all of this technological progress comes a ubiquitous problem known as electronic garbage, or e-waste. The increasing rate of technological obsolescence has made disposing of old devices a major global concern. Conventional waste management techniques are unable to handle the complexity and diversity of electronic components, which calls for advanced techniques for effective e-waste management. The increased amount of e-waste presents numerous risks to the environment and human health, which are compounded by the fact that these gadgets include flame retardants, lead, mercury, and cadmium. The field of e-waste classification in this literature review, delving into the complex array of approaches, innovations, and technologies meant to detect, classify, and recycle electronic trash.

Implementing an information management system for Electrical and Electronic Equipment (EEE) is crucial for efficient e-waste management. Justification stems from observed facts: 1) Government departments and public establishments are major EEE consumers, 2) Government e-procurement systems lack centralization, 3) Inconsistent EEE consumption data across the sector, and 4) Disconnection between consumption data, e-procurement, and disposal. Mandatory coding (indexing) is essential for system implementation. Referring to the UNEP inventory assessment manual, the paper notes constraints in e-waste assessment methodologies. It highlights the significance of EEE consumption data, emphasizing the public sector's substantial e-waste generation. Suggesting an MIS for the public sector underscores the lack of an effective monitoring system despite recent e-waste regulations. Current e-waste determination methods rely on estimations, necessitating standardized record-keeping through a common EEE categorization protocol across organizations. The procurement, consumption, and disposal details need linkage, emphasizing the need for an integrated system updated from the Request for Proposal (RFP) stage to the recycling stage.[18]

The role of smart technologies in collection and classification of e-waste: The main criteria set includes three, which are economic, social and environmental criteria, including a total of 15 sub criteria. Smart technologies identified in this study were robotics, multiagent systems, autonomous tools, smart vehicles, data-driven technologies, Internet of things (IOT), cloud computing and big data analytics. The weights of all criteria were found using fuzzy analytic network process (ANP), and the scores of smart technologies which were useful for collection and classification of e-waste were calculated using fuzzy VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR). The most important criterion was found as collection cost with a weight of 0.167, followed by pollution prevention and control, storage/holding cost and greenhouse gas emissions with weights of 0.115, 0.102, and 0.093, respectively. Autonomous tools were found as the best smart technology for collection and classification of e-waste, followed by robotics and smart vehicles [11].

Green computing, vital for sustainability, explores strategies like virtualization, cloud computing, energy reduction, and e-waste management. Despite a focus on energy minimization, the study underscores the significance of addressing e-waste, with 41.8 million metric tons generated globally in 2014. While literature on energy and cloud computing abounds, research on green computing emphasizing e-waste is lacking. The study aims to explore the feasibility, challenges, implementation, and sustainability of integrating e-waste management into green computing. The lifecycle phases are categorized into Production of EEE, Generation of UEEE, Decision for Fate, and E-waste Processing. The paper advocates solutions like the 3R approach, proper disposal, consumption reduction, virtualization, energy-efficient monitors, and electronic parts testing and reuse. The aim is to highlight the current state of e-waste disposal and emphasize the need for research and development in this domain. [17]

Several approaches have been created, refined, and combined in the quest for e-waste classification with the goal of producing the best results possible.

The method of classifying various e-waste materials using thermal imaging. The created method can be applied to the multi-material classification of recyclables from electronic trash into general groups including glass, metal, plastic, and PCBs. After validating the created method, they were able to classify e-waste with an 80– 89% success rate. [1]

A multi-robot system consisting of a manipulator and conveyor system that can sort construction and demolition debris (CND) has been developed using several techniques for screening and separating e-waste items with different shape, size, density, electric, and magnetic properties. This uses reinforcement learning for gripping tasks and supervised machine learning for material categorization. The primary constraints for visual item sorting are the need for a small workspace and controlled lighting. Not cost effective because the suggested system primarily consists of software, which might lower costs, but more hardware is required. [1]

Waste processing has been automated and streamlined by computer vision technology, resulting in a solid and dependable e-waste management system. It provides a number of advantages that promote longterm advancements in e-waste management techniques. Adoption of this technology will reduce the need for human interaction, processing time, and costs, revolutionizing the way e-waste management is processed. Following a thorough examination of the literature, fifteen enablers are selected and validated by subject matter experts. The causal links are then visualized using an integrated approach of "Interpretative Structural Modelling (ISM)" and "Decision-Making Trial and Evaluating Laboratory (DEMATEL)". An integrated ISM-DEMATEL methodology is used to establish a hierarchy of levels and relationships among the identified enablers. [2]

Laser-induced breakdown spectrometry (LIBS) is used for real-time identification of e-waste polymers, including ABS, PS, PE, PC, PP, and PA. Utilizing emission lines (C, H, N, O) and molecular bands (C2), the study employed KNN and SIMCA classification models to rank polymers. The estimated e-waste polymer content, averaging 21%, varies by device type. Common polymers in e-waste include ABS, PC, HIPS, PP, and PE. Recycling is proposed to integrate polymeric waste into production. LIBS, coupled with discriminant function analysis, identified and classified polymer groups. Classification models (KNN and SIMCA) were developed using the theoretical ratios of the emission lines and molecular bands obtained from LIBS. However, KNN and SIMCA models cannot distinguish between ABS and PS samples, posing a limitation. LIBS advantages include high analytical frequency, portability, and minimal sample preparation, making it a promising tool for e-waste polymer recycling and waste management across industries. [19]

An addition to the Faster R-CNN model for electronic component detection is a computer vision model that incorporates contrastive learning. In order to achieve greater margins of separation between classes, contrastive learning suggests a cost function that considers the idea that members of the same class should be represented as similarly as feasible, while members of different classes should be represented as divergently as possible. [3]

The FICS-PCB dataset consists of 9,912 images of 31 PCBs captured using a DSLR camera and a digital microscope. Annotations for the following six component types are included: integrated circuits, diodes, resistors, inductors, transistors, and capacitors. In contrast to the FSCE [10], the Contrastive Faster RCNN trains the model in a single stage rather than use the TFA (two-stage fine-tuning technique) method. A sizable dataset is available for initial training in a few-shot object detection issue, after which new classes with sparse instances are presented. For this reason, the model in the formulation used for this problem was presented to all classes at the same time. [3]

When applying a few-shot problem method, the contrastive Faster R-CNN outperforms the standard model, yielding an approximately 11% higher mAP. The primary difficulty in the target application—identifying electrical components in PCBs for recycling—lies in the requirement for labelled data and the requirement to perform better than other models after being trained on a sizable dataset. [3]

The goal of this strategy is to take advantage of the widespread adoption and use of smartphones by facilitating information interchange regarding the waste that has to be collected from persons. People would take a picture of the waste item and send it to the server of the garbage collection firm, where it would be automatically identified and categorized. A faster region-based convolutional neural network (R-CNN) was utilized to identify the category and size of the waste equipment in the photos, while a deep learning convolutional neural network (CNN) was used to classify the type of e-waste. The chosen e-waste categories have recognition and classification accuracy ranging from 90 to 97%. [5]

Three common types of household electronics that are thrown away are TVs or monitors, washing machines, and refrigerators. The inputs for the model were these photos. Each object had distinct qualities that were used in the learning process. Information on the three classes of items that the photographs represented was the model's output. The feature-learning module and the classification module are the two primary functional modules of a deep learning CNN. The pooling layers and alternating convolutional layers make up the majority of the feature-learning module. Using m different filters, the convolution layer performed a 2D convolution on the input $h \times w \times c$ image x. A feature map of y of c channels was the result. [5]

The sliding convolutional filters are applied to the input by a 2D convolutional layer. By advancing the filters both horizontally and vertically along the input, the layer convolves the data. Spatial subsampling is mostly handled by the pooling layer, which performs a unique 2D convolution with a stride greater than 1 that typically equals the filter size. Layers in the categorization module are fully coupled. It is a single hidden layer feedforward neural network classifier that makes use of the highest-level feature vector. [5] Although CNN has been very effective at classifying images it requires an extensive amount of labelled training data to achieve optimal performance and such datasets can be time and money consuming to collect and annotate. Traditional CNNs also lack inherent scalability because growing the model's size frequently results in inefficient computation and higher memory needs. It has a limited ability to accommodate global context. [5]

A deep learning method developed using Faster R-CNN with ResNet-50 and ResNet-101 as the network architecture classify waste into three categories. The dataset is comprised of 250 images. The test results applied to the Faster R-CNN ResNet-50 and ResNet-101 methods, it produces an average F1 Score of 63% and 77% respectively [15].

The development of an intelligent waste material classification system, utilizing a 50-layer Residual Network pre-trained (ResNet-50) Convolutional Neural Network (CNN) as an extractor and Support Vector Machine (SVM) for waste classification into categories such as glass, metal, paper, and plastic. This system achieves an accuracy of 87% when tested on a trash image dataset by Gary Thung and Mindy Yang. The combination of CNN and SVM is employed for the recognition and classification of waste. Due to the small size of the trash image dataset, a pre-trained ResNet-50 model is utilized to enhance recognition accuracy. ResNet-50 addresses the vanishing gradient problem in deep CNNs by employing residual modules, allowing for effective learning at earlier layers. The accuracy plateaued at 87% after the 12th epoch during training. The system aims to expedite and enhance the intelligence of waste separation processes, minimizing or eliminating the need for human involvement. [22]

The garbage classification algorithm, based on the ResNet-34 algorithm involves designing the overall garbage bin system, encompassing both the hardware structure and the associated mobile app. Subsequently it undergoes optimization through multi-feature fusion, residual unit feature reuse, and the introduction of a new activation function. The input features are extracted in parallel, and their fusion occurs at the parallel structure's end, employing three parallel routes. The study conducts 200-step training on the MNIST dataset using ResNet-34 and four enhanced algorithms. The proposed classification algorithm's superiority is affirmed through verification with constructed garbage data. Presently, deep learning technology is extensively applied in image classification, with notable achievements such as ResNet's 93.03% and 95.51% accuracy on cifar-10 and Imagenet, respectively, and G.-Q. Zhong et al.'s combined residual and Inception blocks achieving high accuracy on MNIST, SVHN, and cifar-10 datasets. Despite these achievements, there is room for improvement in achieving higher classification accuracy using deep learning algorithms in public datasets. [21]

CNN, YOLO, and faster RCNN-based multi-class classification methods to detect different types of waste. The experimental results for multi-class classification show that CNN has 80% accuracy with 60% of the loss. Whereas the YOLO algorithm shows an accuracy of 88% and a loss of 40%. But the best results were obtained from faster RCNN object detection with API, with an accuracy of 91% and a loss of 16%. [16]

A proposed artificial waste categorization challenge used a convolutional neural network (CNN) to gather and arrange a dataset into seven categories: trash, cardboard, paper, plastic, metal, plastic, glass, and e-waste. then made a distinction between the following transfer learning algorithms: MobilenetV2, EffiecienNetB7, Xception, DenseNet121, and Resnet-50. [4]

By using a technique called transfer learning, a model that was trained on one dataset was then applied to train another dataset that had a different class distribution or even classes that weren't in the training dataset. The model uses unique transfer learning techniques along with categorization algorithms to assist see how the object will change depending on different inputs. dividing the photos into small pieces and then adding them to layers of neural networks, for instance. [4]

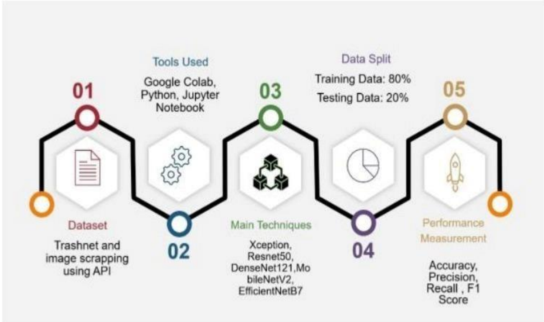

fig. 1. general methodology

Each neural layer needs to shrink the object at that moment while maintaining the intricacies of the image because it employs different filters and algorithms to aid in classifying the object. After accounting for every variable, the model calculates the precision of how likely an item is to fit the trained sample. In order to determine if the features are necessary for the development of the model, this study will examine each function's variable. Subsequently, they have employed data visualization techniques to identify inaccurate photos in the collection and create models that classify and label the images based on the categories. Next, they analyzed the accuracy using performance evaluation as a basis. Ultimately, the optimal algorithm for this task will be demonstrated. [4]

For our model, DenseNet121 obtained a high precision characterization of around 93.3%; Mobilenet likewise showed an amazing 93% conversion to various forms of waste; while Resnet50, Xception, and EfffiecienNetB7 obtained 92%, 92.5%, and 87%, respectively. Other hyperparameter tweaking for the transfer learning methods for future work, developing mobile applications, and integrating real-time object identification techniques like YOLO or related approaches employing microprocessor devices or cameras are areas of attention. These techniques will help increase accuracy. [4]

Computing power is the primary factor in object recognition using a camera; real-time applications on a variety of hardware can be used with many different alternatives. In comparison to CPUs, GPU-based hardware is much more computationally powerful and is frequently employed in that list. When pricing is considered, NVIDIA's Jetson low-cost GPU device is the most often used piece of hardware that makes use of GPUs. Features are automatically extracted from the image in the dataset with the use of deep learning. More datasets are needed for deep learning to achieve higher accuracy. [6]

One of the main uses for transfer learning is the simple training of a deep learning model on a bespoke dataset utilizing a pretrained model that was trained on an already-existing dataset. The transfer learning model's weights are loaded into the current model to give it classification capabilities. Another method for classifying this data is to create a neural network. It is simple to implement a neural network application in hardware with the Jetson Nano development kit. The Jetson Nano has a high frame rate for processing camera data and categorization because it is a GPU enabled device. Deep learning is crucial for e-waste segregation since it does away with the necessity for human processes. This essay focuses on categorizing electronic waste into eight groups: IC voltage regulator, resistor, capacitor, IC voltage regulator, LCD, Relay, Circuit board, Node MCU, Battery. [6]

The development of Focus-RCNet, a lightweight network architecture for automated waste classification designed to overcome the computational challenges of convolutional neural networks in real-time embedded applications. Inspired by MobileNetV2's sandglass structure, the model utilizes deeply separable convolution, SimAM attention mechanism, and knowledge distillation to enhance focus on waste image features while maintaining a small parameter count. The Focus module, employed before the network layer, employs techniques to reduce image size, feature dimensionality, and prevent overfitting. The approach minimizes information loss during down sampling, contributing to the ultimate goal. The model incorporates depth-separable convolution, bottleneck, and inverted residual structures for efficient training. Implemented using PyTorch on an NVIDIA GeForce RTX 3090 Ti server, trained on the TrashNet dataset, and employing data augmentation, the study acknowledges high accuracy but notes limitations in dataset confinement and potential misjudgments. The trade-off between computational complexity and model accuracy remains a challenge, emphasizing the need for further exploration in diverse applications. [20]

Transfer learning is applied by fine-tuning the output layers of AlexNet as a pre-trained model and use dataset that contains 12 classes from 6 smartphone brands. They found that the optimizer of Stochastic Gradient Descent with Momentum and $3e^{-4}$ as a learning rate brings almost 98% model accuracy with generalization. Pros: reducing the error rate of e-waste sorting and it is easier to use transfer learning than building the NN from scratch. Cons: Dataset is very small.[12]

Transformers designed primarily for natural language processing performed exceptionally well on a range of tasks. However, putting transformers directly into images is difficult because of their intrinsic sequential processing nature and lack of spatial inductive biases. This issue was addressed with the Vision Transformer (ViT), which divides a picture into patches and treats them as tokens, enabling transformers to be used for image categorization. ViT has many shortcomings despite its remarkable outcomes. Scalability is limited by its heavy reliance on patches, which results in higher memory consumption. A solution to this problem with CNNs and ViT was the development of data-efficient image transformers (DeiT). The streams are referred to as edge stream for the first and pyramid stream for the second.is Edge Stream: An edge stream in a two-stream DeiT architecture is required because to the significance of obtaining fine-grained edge information in image processing jobs. This stream relies heavily on the sobel operator, which makes it possible to extract edge features from input images. The primary function of the sobel operator is to locate edges by calculating gradients at the locations of each pixel. Prayed Stream, which is necessary for effectively capturing multiscale contextual data. The model's ability to understand and classify objects is enhanced by the pyramid stream, which incorporates data from different receptive field sizes. As a result, the model can handle objects of various sizes and levels of complexity with flexibility, which is crucial for effective visual comprehension tasks. [7]

The focus of this research work is on the machine learning algorithm. They have used an empirical approach to study various machine learning algorithms classifying solid waste, mainly recyclables. Figure 1 shows a simplified approach of the system which uses a camera to capture the image of the object (waste/trash). This image is fed into a trained machine learning algorithm which performs the identification, and a classifier further determines the class and opens the corresponding lid of the bin. Based on Transfer Learning techniques, it is being proposed that a total of 18 different versions of the CNN are considered, with 3 main classifiers, namely SVM, Sigmoid and SoftMax. They identified TrashNet as the dataset to be used for the experiments due to its availability, number of classes and wide range of images. The approach, as depicted in training each CNN algorithm with each classifier and comparing the results using the following parameters: training and testing accuracy, loss and training time.[23]

Waste segregating is the procedure by which waste is isolated into various components. Waste isolation implies separating waste into dry and wet. Dry waste incorporates wood and related items, metals and glass. Wet waste commonly alludes to natural waste for the most part produced by eating foundations and are overwhelming in weight because of sogginess. Waste can likewise be isolated on premise of biodegradable or non-biodegradable waste [24].

The system segregates the waste using CNN algorithm in machine learning. The algorithm detects and classifies the waste according to the dataset provided to CNN. The algorithm classifies the waste as biodegradable and non-biodegradable. As mentioned in the given system uses microcontroller and sensors to segregate the waste directly whereas they look forward to firstly classifying the waste and reduce the use of bigger hardware components. The result is then processed to the hardware components of the system where raspberry pi is being used to dump the waste in the bins. As the system works independently, there is no need of human mediation to control or to do any dreary assignment from this time forward. The system is limited to the objects which look like metals but are not metals.[24]

This proposed model primarily consists of two segments: a separate autonomous RNN component and a CNN component with an image size of (224*224). In the proposed model, an independent, autonomous RNN component has two different LSTM layers, both of which are of batch shape 2048.

However, the CNN is finally carried via a pre-trained TL model, i.e. ResNet-50, InceptionResNet-V2, until it achieves the final destination layer, i.e., Convolutional layer, that includes bottleneck functionalities; these all contain a batch size of 64 Figure presents the working of the proposed Hybrid CNN-LSTM with the TL model.[25]

This proposed model primarily consists of two segments: a separate autonomous RNN component and a CNN component with an image size of (224*224). In the proposed model, an independent, autonomous RNN component has two different LSTM layers, both of which are of batch shape 2048. However, the CNN is finally carried via a pre-trained TL model, i.e. ResNet-50, InceptionResNet-V2, until it achieves the final destination layer, i.e., Convolutional layer, that includes bottleneck functionalities; these all contain a batch size of 64 Figure presents the working of the proposed Hybrid CNN-LSTM with the TL model.[25]

The components' results are combined through a variable size-based multiplication process. The result is procured further into the activation function (Soft-Max) and the categorization layer of different classes for multi-class classifications. In the next phase, all the Convolutional layers with various filter strengths can recognize waste images with magnifying factors because every convolution operation may take images of a specific length. The proposed hybrid model contains the following key components. a) LSTM layer with dense features; b) SoftMax activation function c) fully connected layer component; e) Convolutional layer; f) pooling layer. [25]

In this work, it focuses on using drones for smart waste management. Drones are used for identifying waste using image processing and deep learning and use GPS and GSM methods to send location to authorities. The project has been improved by using trash segregation. This makes garbage handling more effective and convenient for the personnel involved. It considerably reduces labor. The Image Classification Model makes use of PyTorch libraries. Convolutional neural networks (CNNs) are the algorithms used for image categorization. The present model has a 95% accuracy level. Furthermore, the project also involves a multi-object detection model using OpenCV, Darknet, and Yolov3.[26]

The proposed idea is primarily concerned with the identification as well as classification of waste materials dumped in garbage which is depicted in Figure 1. Unsegregated waste is typically dumped in land areas and allowed to the decay process, which can take more than hundreds of years if the dumped waste is a non-biodegradable one. To avoid such issues in the waste segregation process, the proposed work aims to recognize and classify the wastes accurately according to the required categories without any human intervention. The classification of wastes is done merely based on the material of the waste regardless of its shape or size. The learning process of the proposed method is automatic, and it could do the continuous update based on the presence of new materials. The benefits of the proposed method facilitate fast decomposition, reduced health risks, and an easy process that doesn't require any initial investment and it is completely automatic.[26]

The proposed model is more memory efficient and time efficient compared to other existing models and the major drawback is it uses a large dataset [26].

In this study, a new IDRL-RWODC technique has been presented for recycling waste object detection and classification. The IDRL-RWODC technique derives from a Mask RCNN with DenseNet model for the detection and masking of waste objects in the scene. In addition, the DRL based DQLN technique is employed to classify the detected objects into distinct class labels. The detailed working of these processes is given in the succeeding sections.[27]

The IDRL-RWODC technique derives from a Mask RCNN with DenseNet model for the detection and masking of waste objects in the scene. In order to boost the object detection outcomes of the DenseNet model, a hyperparameter optimization using DFA. In addition, the DRL based DQLN technique is employed to classify the detected objects into distinct class labels. The IDRL-RWODC technique has the ability to recognize objects of varying scales and orientations. For ensuring the improved waste classification outcomes of the IDRL-RWODC technique, an extensive experimental analysis takes place to investigate the efficacy in terms of different measures. The experimental results pointed out the better performances of the IDRL-RWODC algorithm over the current techniques [27].

To achieve the objectives of this research, several machine learning models were developed and tested, including traditional as well as deep learning algorithms. This includes Support Vector Machine (SVM), Random Forest (RF), and Deep CNN. These models were compared based on several metrics including F1 score, Precision, Recall and Accuracy. Since it has several basic and effective tools for machine learning, Python programming language is among the most frequently used ones solve diverse problems (Srinath,2017). Therefore, in his study, Python has been utilized in the model's development [28].

The study compares the efficiency and accuracy of CNN models with two other supervised learning algorithms, namely Support Vector Machines (SVM) and Random Forest (RF). To determine the optimum parameters for the traditional model development, Grid Search algorithm was used from the sk-learn interfaces for hyperparameter tuning. For Random Forest (RF), it assumed that the most critical hyperparameters were the number of estimators (n) and the criterion (Gini or entropy). A value of 200 was chosen for the number of trees in the forest (n), and the Gini optimizer was chosen as the tree criterion, as the average training time and cross-validation value are both acceptable. The results showed that the model's accuracy was 72% with a training time of 0.142 minutes and a standard deviation of 0.08. On the other hand, for the SVM model the accuracy was 63% with a training time of 2.43 minutes. Such relatively low accuracies of the traditional models may be attributed to the fact that these models are not suitable for classification of solid waste items from RGB images [28].

The research uses Trashnet dataset that consists of 6 classes of trash objects for the trash image classification. Support vector machines (SVM) with scale-invariant feature transform (SIFT) features and a convolutional neural network (CNN) were used as models. In their experiment, the SVM outperformed the CNN; however, the CNN was not trained to its full potential due to difficulties in determining optimal hyperparameters.[29]

Deep Learning models can be hybridized to improve the accuracy of object classification models. In a study, uses 5000 images with a resolution of 640 by 480 pixels and a plain grey background are used. When the investigated items have strong image features, both the Multilayer Hybrid System and CNN perform well. CNN, on the other hand, performs poorly when waste items lack distinguishing image features, particularly "other" waste. Under two different testing scenarios, MHS achieves significantly higher classification performance: the overall performance accuracies are 98.2 percent and 91.6 percent, respectively.[29]

As trash can belong to different environments, propose a deep learning approach for medical waste identification and classification. The authors propose ResNeXt, a deep learning-based classification method that was applied to 3480 images and successfully identified 8 types of medical waste with an accuracy of 97.2 percent; the average F1-score of five-fold cross-validation was 97.2 percent.[29]

Despite the comparison, there are still insufficient datasets for classifying electronic trash images. As a result, this study provides a comprehensive dataset of Images from eight categories for e-waste classification and describe a two-stream DeiT-based technique for picture classification.

III. Ease of Use

The dataset and the model's architecture are both covered in this section.

3.1 Data Collection and Preprocessing

fig. 2. sample of e-waste vision dataset

The E-Waste Vision Dataset, which consists of 1053 publicly available photos of different electronic equipment, including cameras, laptops, microwaves, mobile phones, televisions, keyboards, and mouse. Figure 2 displays a sample of the dataset. To ensure that the dataset is prepared for effective training and evaluation, many preparation procedures have been employed. To separate the objects of interest, these techniques include rescaling the photos to a typical 348x384 pixel size and adding random rotation, height and weight shift, shear transformation, zooming, and flipping to the dataset. Furthermore, the pre-processed dataset is divided into three sets, each with a ratio of 70:10:20: a training set, a validation set, and a test set.

3.2 Proposed Model

The suggested model features a two streams DeiT architecture with real-time classification and audio output, as seen in Figure 3. The streams are referred to as edge stream for the first and pyramid stream for the second.

1) Edge Stream: In a two-stream DeiT architecture, an edge stream is required because to the significance of obtaining fine-grained edge information in image processing tasks. In this stream, the Sobel operator is essential because it makes it possible to extract edge features from input images. The primary function of the Sobel operator is to locate edges by calculating gradients at the locations of each pixel.

To match the input dimension of the pre-trained DeiT, a convolution layer is utilized in the edge stream after the Sobel operator. In this layer, a collection of learnt filters are used to convolve the Sobel outputs. Finally, the stream introduces the DeiT model, which makes use of self-attention mechanisms to record contextual linkages among edge features as well as global context.

2) Prayed Stream: To effectively capture multiscale contextual information in a two-stream DeiT architecture, a pyramid stream is necessary. By incorporating elements from, the pyramid stream enhances the model's ability to recognize and classify items across different receptive field widths. This allows the model to be flexible enough to handle objects with varying sizes and degrees of complexity, which is essential for tasks involving visual comprehension.

The pyramid stream relies on the ASPP block, which gathers multi-scale contextual data. It does this by utilizing parallel convolutional layers with different dilation rates. Convolutional filters can gather characteristics at various scales since the dilation rates regulate the size of their effective receptive fields. Our model's ASPP block consists of five convolution layers, each of which has a different dilation rate (1, 2, 3, 4, and 5). These layers are accompanied by fewer filters (64, 32, 16, 8, and 4, respectively) to guarantee computational efficiency.

The need for CBAM arises from the need to obtain attention both spatially and channel-wise. Through adaptive feature map recalibrating, the CBAM module increases the model's selectivity [25]. It uses a two-step process to accomplish this, as shown in Figure 5. First, the spatial attention mechanism records fine-grained spatial dependencies by concentrating on visually relevant locations. Second, the channel attention mechanism learns to highlight relevant channels, which facilitates the extraction of discriminative features.

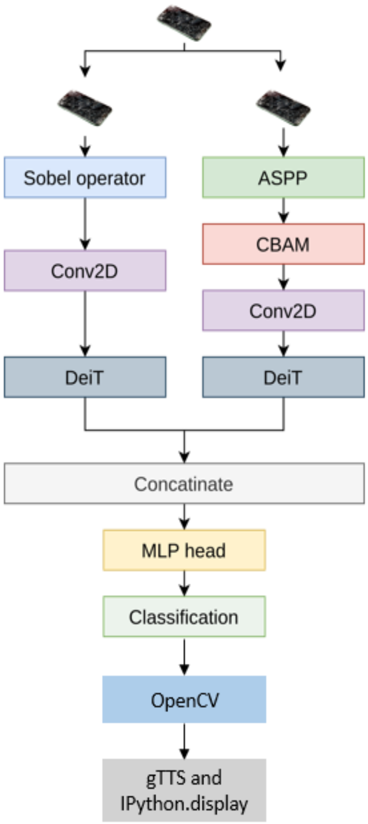

fig. 3. proposed architecture

This dual attention approach effectively captures both local and global contextual information, which helps the model distinguish between different object types. As a result, the ASPP block is integrated before the CBAM module. A convolution layer is once more applied after the ASPP and CBAM blocks in order to match the input dimension of DeiT. Additionally, this layer aids in the fusion of multi-scale contextual data from the ASPP block. Ultimately, the DeiT model that has been trained is included into the stream.

3) Both Streams Concatenate: In our research, the unique qualities of the two streams are combined to enhance the overall performance of categorization. These attributes are taken from the edge stream and the pyramid stream, respectively, and then integrated to form a single representation. After concatenating, this feature vector is fed into a Multi-Layer Perceptron (MLP) for further processing and categorization.

The MLP design consists of several fully connected (FC) layers with various configurations. An FC layer with 512 neurons is followed by a dropout layer with a 30% dropout rate. By removing neurons at random during training, this dropout layer helps avoid overfitting. Subsequently, a 20% dropout layer and a fully connected layer with 256 neurons are added. In order to reduce the number of dimensions in the feature representation, a bottleneck layer consisting of 256 dense neurons is added last. Each FC

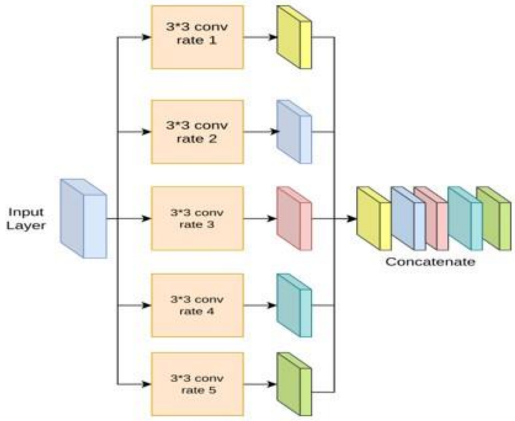

fig. 4. architecture of the ASPP block

layer in the MLP is activated using the Rectified Linear Unit (ReLU) activation function. ReLU adds non-linearity to the model, enabling it to pick up complex patterns and representations. The last layer of categorization is SoftMax layer.

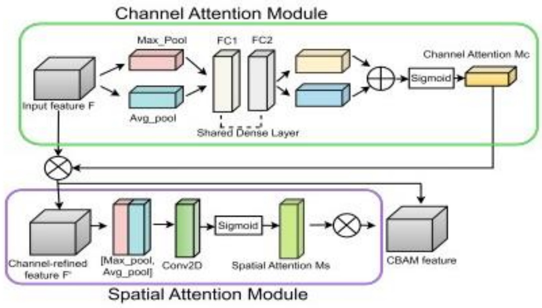

fig. 5. convolutional block attention module for proposed model

The whole model is returned by the model construction function, which accepts the picture height and weight as input. The pyramid stream is created by the function once the edge stream is first created. In the combine part, they are later combined. Ultimately, the model is given back with two input layers from two steams and an output layer that is the MLP head's softmax layer.

The computational efficiency of the fusion approach is one noteworthy aspect. Using only 20 training epochs and less than 1 million trainable parameters, the model finds a compromise between efficacy and complexity. By utilizing the concatenated features and the MLP architecture, the fusion approach effectively combines the complementing data from the edge and pyramid streams, enhancing classification performance while keeping a reasonable number of parameters.

3.3 Audio Output

For audio output in our integrated system, audio processing technologies such as gTTS (Google Text-to-Speech) have been utilized. gTTS (Google Text-to-Speech) is a Python library and CLI tool that interfaces with Google's Text-to-Speech API. It allows you to convert text into speech in various languages and voices. gTTS can generate audio files directly from text strings, and it supports several output formats, including mp3 and wav.

Key features of gTTS: Simple Interface: gTTS provides a simple Python API for converting text to speech with just a few lines of code. Language Support: It supports multiple languages, including English, French, Spanish, German, Italian, and many more. Voice Selection: You can select different voices for the speech synthesis, including male and female voices with different accents and characteristics. Speed and Pitch Control: gTTS allows you to adjust the speed and pitch of the generated speech. Output Formats: It supports multiple output formats, such as mp3, wav, and ogg. Free and Online: gTTS utilizes Google's Text-to-Speech

API, which is free to use within certain limits. However, you'll need an internet connection to use gTTS as it relies on Google's cloud services.

Generate Speech from Predictions: Convert the model's output predictions into speech using a text-to-speech (TTS) synthesis library. This would involve converting the output predictions (which are typically in a numerical or categorical format) into human-readable text, and then using a TTS system to generate speech from that text.

Integrate TTS System: Integrate a TTS system into your codebase. There are various TTS libraries available in different programming languages that you can use for this purpose. Some popular options include Google Text-to-Speech (gTTS) for Python, Microsoft Azure Speech Service, and Mozilla TTS.

Call TTS System: After obtaining the model predictions, call the TTS system with the generated text to produce the corresponding audio output.

The IPython.display module is a powerful component of IPython, a popular interactive computing environment for Python. It provides functionalities to display various types of rich media content, including images, audio, video, HTML, and interactive widgets, within IPython environments such as Jupyter Notebooks. It allows you to embed images, multimedia, plots, and formatted text directly into your code. The display.Audio class is used to play audio files within IPython environments. It supports various audio formats such as WAV, MP3, OGG, etc., and allows users to control playback options such as volume, autoplay, and loop.

IV. Results and Discussion

The results and discussion that follow are the first two sections of this section, which starts with a brief explanation of the evaluation metrics used to evaluate the model.

4.1 Evaluation Matrix

Our model for categorizing e-waste was evaluated using several parameters such as accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC). These measurements demonstrate how to identify each class while balancing precision and recall, providing insights on overall classification accuracy. To better understand the model's performance, refer figure 6 and figure 7 for the confusion matrix and receiver operating characteristic (ROC) curve respectively.

The confusion matrix indicates how accurately the model classified each class. Based on the analysis in the confusion matrix, it is found that four instances of the mobile class were incorrectly classified as TV, one instance of the microwave class as TV, one instance of the mobile class as a smartwatch, and one instance of the camera class as a smartwatch. Despite the classification failures, the model accurately predicted the remaining occurrences across all classes.

The ROC curve illustrates the trade-off between TPR and FPR at different categorization criteria. Our ROC curve, in the test dataset, all classes achieved ROC scores of 1, indicating excellent accuracy in distinguishing between positive and negative cases. The micro-average ROC score, which measures overall performance across classes, was the highest attainable value of 1. Our program accurately and efficiently classifies e-waste photos, as evidenced by evaluation metrics and outcomes. Our model is reliable and suitable for e-waste management because of its excellent accuracy, precision, recall, F1 score, and MCC values.

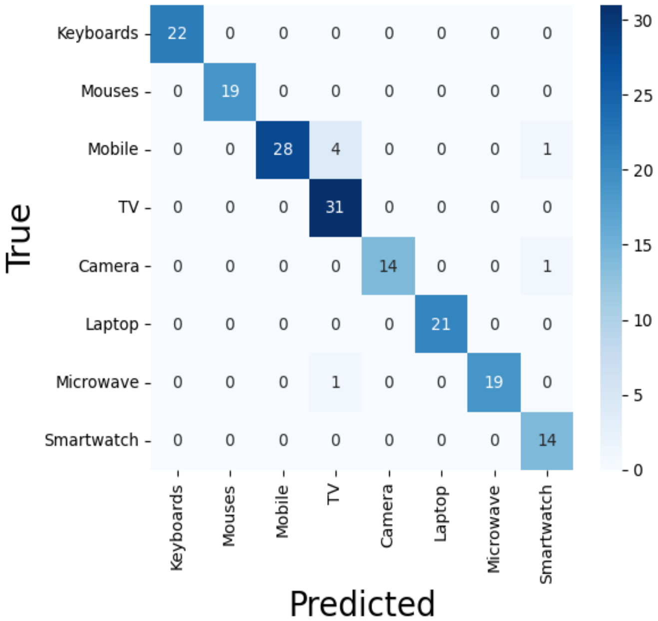

fig. 6. confusion matrix

4.2 Validation Results

Our model achieved an accuracy of 0.93 in the validation set. Precision measures the proportion of appropriately identified instances among predicted positive examples. The score was 0.93. Similarly, the recall score of 0.92 indicates that the majority of positive cases were accurately recognized. The precision and recall-balancing F1-score was 0.92. The MCC, which assesses classification accuracy by considering true and false positives and negatives, yielded a value of 0.91.

4.3 Test Results

The model's accuracy for the test set was 0.96, indicating its robustness and generalizability. The precision score of 0.96 indicates great accuracy. Percentage of accurately classified incidents. The recall score of 0.9670 indicates that the model accurately identifies positive cases. The F1 score was 0.96. The test sets MCC score of 0.95 indicates a strong connection between expected and actual labels.

The integration of gTTS for audio output and IPython.display for displaying audio files enhances the interpretability and accessibility of the classification results, facilitating easier understanding and utilization of the model's predictions.

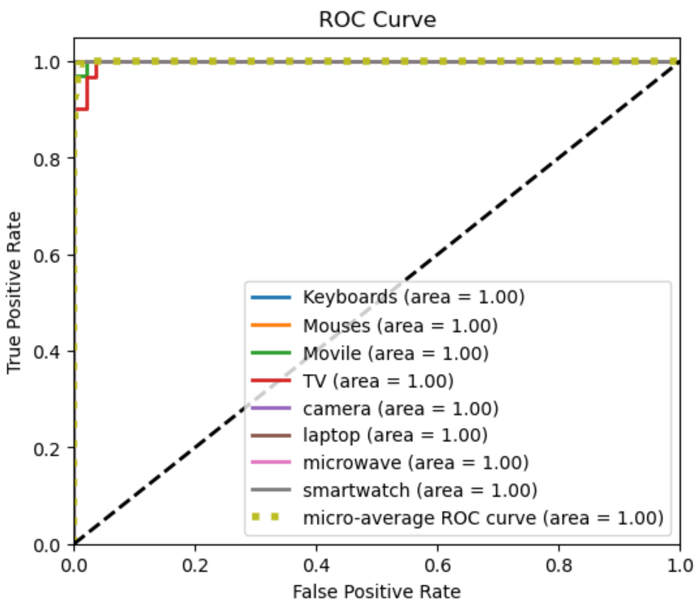

fig. 7. receiver operating characteristic curve

V. CONCLUSION

In conclusion, our research underscores the critical need for innovative and interdisciplinary approaches in tackling the growing challenge of electronic waste (e-waste). The importance of robust classification techniques, spanning from thermal imaging to advanced deep learning models like the two-stream DeiT approach has been emphasized. Additionally, our study highlights the integration of real-time processing and audio output capabilities using computer vision technology, further enhancing the adaptability and utility of our proposed solution.

By harnessing emerging technologies and fostering synergies between diverse methodologies, significant potential to develop comprehensive and adaptable systems for effectively managing the intricate and hazardous composition of e-waste can be unlocked. These systems not only offer promise in mitigating the adverse environmental and public health impacts associated with e-waste but also promote sustainable waste management practices on a global scale.

As technology continues to evolve, continuous research and innovation are imperative to address the evolving complexities of e-waste. Through collaborative efforts and ongoing advancements in technology, the way towards a more sustainable and environmentally conscious future is good.

REFERENCES

[1] G. Sathish Paulraj1;H. Subrata, Ph.D., M.ASCE2; T. Amber3; and T. Atul, Ph.D. "Classification of Recyclables from E-Waste Stream Using Thermal Imaging-Based Technique", ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, December, 2018, 69

[2] Himansu sharma, harish kumar, Sachin mangla, "ENABLERS TO COMPUTER VISION TECHNOLOGY FOR SUSTAINABLE E WASTE MANAGEMENT", International Journal on Perceptive and Cognitive Computing (IJPCC), Volume 412, August, 2023

[3] Leandro H. de S. Silva1,2 Agostinho Freire1 Bruno J. T. Fernandes1 George O. A. Azevedo1 and Sergio C. Oliveira1."Evaluation of Contrastive Learning for Electronic Component Detection", ESANN 2023 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 4-6 October 2023, i6doc.com publ., ISBN 978-2-87587-088-9, 489

[4] Zian Md Afique Amin, Khan Nasik Sami, Raini Hassan. Journal of Cleaner production, "An Approach of Classifying Waste Using Transfer Learning Method", International Journal on Perceptive and Cognitive Computing (IJPCC) Vol 7, Issue 1 (2021), 41

[5] Piotr Nowakowski * , Teresa Pamuła , "Application of deep learning object classifier to improve e-waste collection planning", Journal of Physics: Conference Series, Volume 1997, Asian Conference on Intelligent Computing and Data Sciences (ACIDS), April, 2020.

[6] S. Elangovan, S. Sasikala, S. Arun Kumar, M. Bharathi, E. Naveen Sangath, and T. Subashini, "A Deep Learning Based Multiclass Segregation of E-waste using Hardware Software Co-Simulation", Journal of Waste Management 109:1-9, 2021 J. Phys.: Conf. Ser. 1997 012039., 1.

[7] Niful Islam, Md. Mehedi Hasan Jony, Emam Hasan, Sunny Sutradhar, Atikur Rahman, Md. Motaharul Islam , "EWasteNet: A Two-Stream Data Efficient Image Transformer Approach for EWaste Classification", September , 2023

[8] Baldé, C. P., A., Forti, V., Garbarino, S., & Steger, C. (2023). The Global E-waste Monitor 2023. United Nations University (UNU)/International Telecommunication Union (ITU) / International Solid Waste Association (ISWA).

[9] M.D.K.C.Jayathilakaa*, T.Manori Pereraa, 2024, Suitability of E-Waste as an Alternative for Coarse Aggregate in Concrete, Prof. Nilmini Liyanage, Dr. Dhanushka Udayanga Dr Akila Wijethunge, International Conference on Innovation and Emerging Technologies 2023 (ICIET2023), University of Sri Jayewardenepura, Sri Lanka, November 23rd-24th, 2023.

[10] Umang Khandelwal1,a, Sonali Bondre1, Mayur Shirish Jain2, Kunwar RaghavendraSingh3, 2023, Present scenario of E-waste generation, legislation,management and impacts: A review in Indian context, Prof. Sudhir Kumar Goyal, Prof. Subhash Chnadra Tripathi , 2nd International Conference on Futuristic and Sustainable Aspects in Engineering and Technology(FSAET-2021), ¬ GLA university Mathura, India, December 24th-26th, 2021.
[11] AV Shreyas Madhav, Raghav Rajaraman, S Harini, and Cinu C Kiliroor, Application of artificial intelligence to enhance collection of E-waste: A potential solution for household WEEE collection and segregation in India, Waste Management & Research 2022, Vol. 40(7), November,2021, 1047–1053.
[12] Nermeen Abou Baker1, *, Paul Szabo-Müller1 and Uwe Handmann1, Transfer learning-based method for automated e-waste recycling in smart cities, EAI Endorsed Trans Smart Cities, vol. 5, no. 16, p. e1, Apr. 2021.
[13] Dr. Govind Singh and Jitendra Singh Bhadauria. Statistical modeling of electronic waste (e-waste) generated in India. International Journal of Statistics and Applied Mathematics. 8. 10.22271/maths.2023.v8.i6Sb.1368. Volume 8, November, 2023,116-122.
[14] Lobo E, D'souza CV, Lasrado SA Study and Analysis of Disposal of Electronic Waste and Its Effect On the Environment. International Journal of Latest Trends in Engineering and Technology Special Issue SACAIM;c2016; p. 273-277.
[15] Ma'rifah, Puteri & Sarosa, Moechammad & Erfan, ROHADI. (2023). Comparison of Faster R-CNN ResNet-50 and ResNet-101 Methods for Recycling Waste Detection. International Journal of Computer Applications Technology and Research. 12. 26-32. 10.7753/IJCATR1212.1006. Volume 12–Issue 12, December, 2023, 26-32.
[16] Narayanswamy, Neeraja & Rajak, Abdul & Hasan, Shazia. (2022). Development of Computer Vision Algorithms for Multi-class Waste Segregation and Their Analysis. Emerging Science Journal. 10.28991/ESJ-2022-06-03-015. Volume 6, April, 2022, 631-646.
[17] Biswajit Debnath, Reshma Roychoudhuri, Sadhan K. Ghosh, E-Waste Management. A Potential Route to Green Computing,Procedia Environmental Sciences, Volume 35, 2016, Pages 669-675,ISSN 1878-0296
[18] Sashi Kumar, Shatrunjay Rawat, Future e-Waste: Standardisation for reliable assessment, Government Information Quarterly, Volume 35, Issue 4, Supplement, 2018, Pages S33-S42, ISSN 0740-624X
[19] Vinicius Câmara Costa, Francisco Wendel Batista Aquino, Caio Marcio Paranhos, Edenir Rodrigues Pereira-Filho, Identification and classification of polymer e-waste using laser-induced breakdown spectroscopy (LIBS) and chemometric tools, Polymer Testing, Volume 59, 2017, Pages 390-395, ISSN 0142-9418
[20] Zheng, D., Wang, R., Duan, Y. et al. Focus-RCNet: a lightweight recyclable waste classification algorithm based on focus and knowledge distillation. Vis. Comput. Ind. Biomed. Art 6, 19 (2023)
[21] Z. Kang, J. Yang, G. Li and Z. Zhang, "An Automatic Garbage Classification System Based on Deep Learning," in IEEE Access, vol. 8, pp. 140019-140029, 2020
[22] Olugboja Adedeji, Zenghui Wang, Intelligent Waste Classification System Using Deep Learning Convolutional Neural Network, Procedia Manufacturing, Volume 35, 2019, Pages 607-612, ISSN 2351-9789
[23] Nadish Ramsurrun1,GeerishSuddul2,SandhyaArmoogum,Ravi Foogooa, Recyclable Waste Classification Using Computer Vision And Deep Learning, vol.20x, 2022-june
[24] Yesha Desai1,Asmita Dalvi, Pruthviraj Jadhav , Abhilasha Baphna,International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887 Volume 6 Issue III, March 2018
[25] Umesh Kumar Lilhore,Sarita Simaiya,Surjeet Dalal • Robertas Damaševičius, A smart waste classifcation model using hybrid CNN LSTM with transfer learning for sustainable environment, Revised: 24 August 2023 / Accepted: 27 August 2023
[26] Gayathri Rajakumaran,,Shola Usharani,,Christie Vincent, Sujatha M, Smart Waste Management: Waste Segregation using Machine Learning, doi:10.1088/1742-6596/2471/1/012030,2023
[27] Mesfer Al Duhayyim,Taiseer Abdalla Elfadil Eisa, Fahd N. Al-Wesabi, Abdelzahir Abdelmaboud , Manar Ahmed Hamza, Abu Sarwar Zamani, Mohammed Rizwanullah6 and Radwa Marzouk, Deep Reinforcement Learning Enabled Smart City Recycling Waste Object Classification, vol.71, no.3,Received: 17 October 2021; Accepted: 29 November 2021

[28] Hani Abu-Qdais, Nawras Shatnawi, Esra'a AL-Alamie, Intelligent solid waste classification system using combination of image processing and machine learning models, February 14th, 2023

[29] Haruna abdu, Mohd halim, Mohd noor,A survey on waste detection and classification using deep learning,publisher-IEEE, Volume: 10, December 2022, 128151 – 128165