

A Deep Learning Approach to Manage and Reduce Plastic Waste in the Oceans.

Abdellah El zaar^{1,*}, Ayoub Aoulalay^{2,**}, Nabil Benaya^{1,***}, Abderrahim El mhouti^{3,****}, Mohammed Massar^{2,†}, and Abderrahim El allati^{1,‡}

¹Laboratory of R&D in Engineering Sciences, FST Al-Hoceima, Abdelmalek Essaadi University, Tetouan, Morocco.

²PMIC Laboratory, FST Al-Hoceima, Abdelmalek Essaadi University, Tetouan, Morocco.

³Faculty of science tetouan, Abdelmalek Essaadi University, Tetouan, Morocco.

Abstract. The accumulation of plastic objects in the Earth's environment will adversely affect wildlife, wildlife habitat, and humans. The huge amount of unrecycled plastic ends up in landfill and thrown into unregulated dump sites. In many cases, specifically in the developing countries, plastic waste is thrown into rivers, streams and oceans. In this work, we employed the power of deep learning techniques in image processing and classification to recognize plastic waste. Our work aims to identify plastic texture and plastic objects in images in order to reduce plastic waste in the oceans, and facilitate waste management. For this, we use transfer learning in two ways: in the first one, a pre-trained CNN model on ImageNet is used as a feature extractor, then an SVM classifier for classification, the second strategy is based on fine tuning the pre-trained CNN model. Our approach was trained and tested using two (02) challenging datasets one is a texture recognition dataset and the other is for object detection, and achieves very satisfactory results using two (02) deep learning strategies.

key-words: Plastic wast recognition, plastic texture recognition, Deep learning, Convolutional Neural Network (CNN), Support Vector Machine (SVM).

1 Introduction

Plastic waste has become one of the most dangerous issues that threatens the stability of life on earth. 88% of the sea's surface is polluted by plastic waste, which is harming animal and human health. In more developed countries, plastic pollution is recovered and recycled [1]. However, in the developing world, garbage collection systems are often inefficient or nonexistent, and generated plastic waste is carried out to sea, and ends up in the water. For this reason we used deep learning to recognize plastic texture in order to detect the plastic waste and pollution in the environment and especially in the ocean. Computer vision and deep learning will facilitate recovering plastic waste from the oceans and protect the wildlife. Therefore, we investigated in Deep learning algorithms to reduce the effects of the plastic pollution crisis. Using such intelligent systems to recognize diverse plastic waste texture is a significant challenge for researchers.

Promoting the waste management and encouraging recycling at the moment is not enough the correct the ecological damage caused by the very large quantity of plastic waste. This crisis needs an effective and fast solution to save the wildlife. The use of Artificial Intelligence as a

complementary will be the most sustainable way to reduce pollution, especially in deep oceans and places that people cannot reach.

In the last years, deep learning and image processing applications have attained a great success in many fields[2]. However, texture recognition using deep learning still an active area of research. The diversity of material textures and features makes the task challenging. For this reason, we investigate in Convolutional Neural Network (CNN) to recognize plastic texture in images. Convolutional Neural Network (CNN) algorithm revolutionized the field of objects recognition and image processing, giving its powerful capacity to work with grid structured data [3]. The main difference between the traditional Feed Forward Neural Network (FFNN) and (CNN), is that the connection between the layers of FCNN is fully connected, while the connection in CNN is organized and designed carefully. Traditional machine learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors and Neural Network algorithms shows better performance for lesser amounts of input data. In contrast, the performance of deep learning algorithms increased with respect to the increment in the amount of data. But in this paper, we will show that deep learning algorithms can also work and provide satisfactory performance in case of small datasets. To perform the recognition process, we used two different learning strategies using Convolutional Neural Network (CNN). The first one is feature extraction using CNN. The convolutional blocks of a pre-trained

*e-mail: abdellah.elzaar@gmail.com

**e-mail: aoulalayayoub@gmail.com

***e-mail: nabil.benaya@gmail.com

****e-mail: abderahim.elmhouti@gmail.com

†e-mail: massarmed@hotmail.com

‡e-mail: abdou.allati@gmail.com

CNN on a large-scale image dataset are used to extract features from the input image, these features are then used to train an SVM classifier. The second strategy is fine-tuning, in this strategy, we replace the fully connected layers of the pre-trained model with a randomly initialized layer. We use a VGG16 model pre-trained on ImageNet [4].

We trained and tested our proposed approaches using two different datasets : a wide material texture dataset that represents well the variety of textures and classes, and a small object recognition dataset.

The rest of the paper is organized as follows. In the first section, we give a brief overview of some significant recent contributions to materials texture recognition. Section 3 presents our proposed approach and Section 4 describes the experiments conducted and compares their results with other leading works. Finally, Section 5 concludes the paper and presents some future research perspectives.

2 Related work

Material texture recognition is a challenging issue due to the diversity and similarity of texture materials. The detection of plastic in the sea is classified as the detection of texture materials of an object in nature. For computer vision tasks, machine learning and deep learning techniques are widely used due to their high performance in tasks such as object detection, scene segmentation, character recognition, etc. Material recognition using machine learning and deep learning techniques is a well-studied topic. We found that several research studies have also been conducted in the area of classification and recognition of materials in the wild. Many studies have focused on hand-engine features to describe the material that makes up an object, In the work [5], the authors used a combination of color and texture features to describe the materials of five Moroccan decorative patterns: glass, wood, fabric, ceramic and plaster. The extracted features are used to train an SVM and KNN classifiers, and they achieve 87% classification accuracy. In their work [6], the authors developed a material recognition system based on the combination of a set of low-level and medium-level features. Dictionaries are concatenated from this set of features and the optimal combination of these features is learned using Dirichlet latent allocation LDA, this technique is based on the assumption that it is not known exactly what the best features are to describe materials. The work [7] used the same method to describe the material by combining a set of features but instead of LDA they use SVM classifier and they improve the classification results by 8.5%. The disadvantage of using craft characteristics to describe materials is that we don't know exactly what the optimal descriptor is, materials often have the same color, texture and shape. A proposed solution to that problem is by using a deep convolutional neural network (DCNN) to describe the material of the image where the optimal description of the image is learned by backpropagation. In the work [8], the authors used CNN to extract features from a material dataset

and then used ensemble learning of knowledge-based classifiers to classify the materials that make up the image, they achieved a mean accuracy of 85%. In their work [9], the authors proposed a transfer learning-based approach to classify the "FMD" material dataset, they used the knowledge from LeNET by copying the first convolutional layers while fine-tuning the other remaining layers. The authors of [10] proposed a deep learning method based on the inception-v3 model to classify waste images, and they achieved 92.5% accuracy.

3 Proposed approach

In our work, we investigate in the modern CNNs that are built using the same principles as traditional ones. CNN algorithm is powerful, and provides effective and accurate results with grid structured data like images. Unlike traditional machine learning algorithms such as (SVM, KNN, etc.), CNN gives better performance using wide datasets. In our study, we will demonstrate that CNN can work also with small databases when using transfer learning strategies. For that reason we worked with two small and challenging datasets, one for waste texture and the other one for waste objects. In Convolutional Neural Networks the layers architecture is organized carefully. CNNs use alternating convolution layers (CL) and max-pooling layers (MP) followed by fully connected layers (FC). The convolution operation is a dot product between an input layer and the filter. The filters(Kernels) are three dimensional parameters $Z_n \times Z_n \times d_n$ which represents the network parameters. The dimensions (length and width) of the output layer after performing the convolutional operation is :

$$L(n+1) = L_n - Z_n + 1 \quad (1)$$

$$W(n+1) = W_n - Z_n + 1 \quad (2)$$

The convolutional operation from the n th layer to the $n+1$ is defined as follows :

$$M_{ijz}^{n+1} = \sum_{r=1}^{Z_n} \sum_{s=1}^{Z_n} \sum_{k=1}^{d_n} w_{rsk}^{(z,n)} h_{i+r-1, j+s-1, k}^{(n)} \quad (3)$$

The most common known CNN architectures, we mention : AlexNet (Krizhevsky et al., 2012)[11], ZF Net (Zeiler & Fergus, 2014)[12], GoogLeNet (Szegedy et al., 2015)[13], VGGNet (Simonyan & Zisserman, 2014)[14] and ResNet (He, Zhang, Ren, & Sun, 2016)[15]. The Figure 1 illustrates performances achieved by the mentioned CNN architectures. It can be seen that each architecture beats the performance of its predecessor, achieving the best accuracy in 2015 with a deep ResNet.

Our model architecture is built around the VGGNet CNN and use both transfer learning and Combined CNN-SVM strategies, without pre-processing step. The output shape of the image data after every single layer is detailed in Table 1.

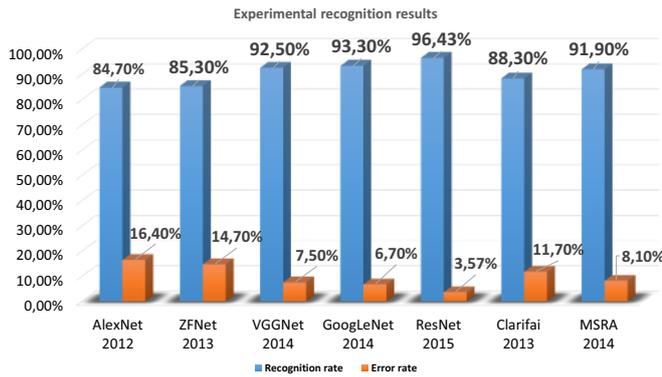


Figure 1: State-of-the-art CNN Architectures and their performances in term of loss and accuracy.

Table 1: The output shape of the image data after every single layer

Layer (type)	Output Shape
input-1 (InputLayer)	(None, 93, 93, 3)
block1-conv1 (Conv2D)	(None, 93, 93, 64)
block1-conv2 (Conv2D)	(None, 93, 93, 64)
block1-pool (MaxPooling2D)	(None, 46, 46, 64)
block2-conv1 (Conv2D)	(None, 46, 46, 128)
block2-conv2 (Conv2D)	(None, 46, 46, 128)
block2-pool (MaxPooling2D)	(None, 23, 23, 128)
block3-conv1 (Conv2D)	(None, 23, 23, 256)
block3-conv2 (Conv2D)	(None, 23, 23, 256)
block3-conv3 (Conv2D)	(None, 23, 23, 256)
block3-pool (MaxPooling2D)	(None, 11, 11, 256)
block4-conv1 (Conv2D)	(None, 11, 11, 512)
block4-conv2 (Conv2D)	(None, 11, 11, 512)
block4-conv3 (Conv2D)	(None, 11, 11, 512)
block4-pool (MaxPooling2D)	(None, 5, 5, 512)
block5-conv1 (Conv2D)	(None, 5, 5, 512)
block5-conv2 (Conv2D)	(None, 5, 5, 512)
block5-conv3 (Conv2D)	(None, 5, 5, 512)
block5-pool (MaxPooling2D)	(None, 2, 2, 512)
batch-normalization (BatchNo)	(None, 2, 2, 512)
dropout (Dropout)	(None, 2, 2, 512)
flatten (Flatten)	(None, 2048)
fc1 (Dense)	(None, 256)
batch-normalization-1 (Batch)	((None, 256)
dropout-1 (Dropout)	(None, 256)
predictions (Dense)	(None, 2)

3.1 Strategy-1- Transfer learning CNN

Transfer learning [16, 17] is the process of transferring knowledge from a source model trained on a source dataset to another target model with a target dataset and task. For image classification, transfer learning is widely used to solve the problem of small training datasets. In this work, the convolutional layers of the pre-trained model (ImageNet) are transferred to our model (VGGNet)

to classify our waste images. We use fine tuning of the pre-trained VGG16 model on ImageNet. The first convolutional layers of a pre-trained VGG16 model are used to extract general discriminative features. The final convolutional layers that extract high level features are initialized with the weights of the pre-trained network and are updated during the training of the network to be able to extract specific features about our dataset. A new randomly initialized fully connected layer is placed in the output of the convolutional part to classify our dataset.

3.2 Strategy-2- CNN-SVM : CNN as feature extractor

Support Vector Machine classifier(SVM) works by representing the multiple classes of a dataset in a space. These classes of data are then separated by a line which called hyperplane. The strength of the SVM algorithm lies in its ability to find the optimal hyperplane that separates between classes. The SVM algorithm is limited to work with large amount and noisy data. In our work, the convolutional layers of the pre-trained VGGNet model are used to extract features from our datasets these features are then used to train the SVM classifier, we believe that a VGGNet model trained on ImageNet has learned a very good representation of the image, and that the learned weights can be used in other tasks such as waste classification.

4 Experiments, comparison and discussion

4.1 Data

The dataset is the challenging part in the recognition process. The recognition system needs a database that represents well the multiple variations of styles and texture features. Large material texture databases are confidential and not available for researchers. The few available ones contain noisy images and non organized classes. For that reason, and in order to improve our proposed method, we have used two datasets: A collected dataset from google images PLAWO-40 and trashNet datasets. In this section, we describe theses datasets.

4.1.1 PLAWO-40 dataset

PLAWO-40 is a dataset of material texture which we recently built. The dataset contains 40 images collected divided into two (02) classes : plastic bags material and wood textures. The dataset can be accessed through this link : <https://www.kaggle.com/abdellahelzaar/plastic-and-wood-pollution-texture-dataset>.

4.1.2 trashnet dataset

Trashnet is a dataset proposed by M. Yang et al (2016) [18], The dataset contains images of recycled objects across six

classes with about 500 photos each. The dataset was hand collected and contains about 99% of all recycled material. In our case we trained our system to recognise two classes of trashnet dataset (Plastic bottles and metal) Figure 2 these two classes presents the majority of dangerous waste in deep sea.



(a) plastic trash



(b) metal trash

Figure 2: trashnet dataset

The PLAWO-40 dataset is a texture recognition dataset, it contains the texture of two objects (wood and plastic), while the second dataset (trashnet dataset) is an object recognition dataset. The reason why we use these two datasets is that we want to evaluate the performance of our models in different scenarios.

4.2 Experiments and results

The CNN algorithm was implemented using Python language. We used several libraries that contains visualisation and image processing functions such as *TensorflowTM*, *KerasTM* Models and *SklearnTM*. We run our CNN model on Microsoft Azure virtual machine with six (06) core processor and 56Go of RAM. Working with Convolutional Neural Networks with a high number of epoches can be time-consuming process, For this reason we used TESLA K80 NVIDIA GPU. GPU boost gives superior performance to our Model and accelerates Libraries and training process.

4.2.1 Experiments on PLAWO-40 dataset

First, we randomly divide our dataset into two parts: 80% for training and the remaining 20% for testing to avoid the well-known phenomenon of over-fitting. The hyper-parameter technique is used for a dropout optimization in both the convolutional layers and fully connected layers.

A Mini Batch size of 32 images is selected, the number of epochs for each method was: 100 epochs, initial learning rate was: 0.001. The metric used to evaluate the performance of our model is the accuracy. This metric consists of dividing the number of correct predictions by the total number of predictions.

Table 2: Achieved recognition rate using PLAWO-40 dataset

Method	with pre-processing Accuracy(%)	without pre-processing Accuracy(%)
Transfer learning using CNN	99%	98.46%
Combined CNN-SVM	98%	95%

It can be seen from Table 2 that using transfer learning CNN with pre-processing achieves better recognition rate with 99% compared to same strategy without pre-processing step, which obtains 98.46%.

With the combined CNN-SVM strategy, we can hold an accuracy of 98%, better than the obtained accuracy of 95% without using pre-processed images. That can be explained by the importance of the pre-processing operation in traditional Machine Learning algorithms such as SVM. The transfer learning strategy outperforms the combined CNN-SVM strategy, we get perfect recognition rate with all instances correctly classified. This shows the power of such technique to work with small datasets and recognise texture images. To evaluate and test the performance of our model, we present the confusion matrix of the learning strategies using pre-processing.

It can be observed from Figure 3 that the values in the diagonal of the confusion matrix are higher than others, these values represents the number of well recognized images in the test-set. This improves the high test accuracy that we have obtained in Table 2.

4.2.2 Experiments on trashnet dataset

We repeated the same set of experiments using our method but on trashnet dataset. We performed the CNN model with two (02) strategies : Transfer learning CNN and combined CNN-SVM. The modified trashnet trainset contains 592 images of metal and plastic objects and 300 images in testset. The Table 3 illustrates the accuracies of our proposed method, all providing good trash objects classification.

Table 3: Achieved recognition rate using trashnet dataset

Method	with pre-processing Accuracy(%)	without pre-processing Accuracy(%)
Transfer learning using CNN	99.58%	98%
Combined CNN-SVM	97%	94%

It can be seen that the use of transfer learning CNN with pre-processing on trashnet dataset gives a better result

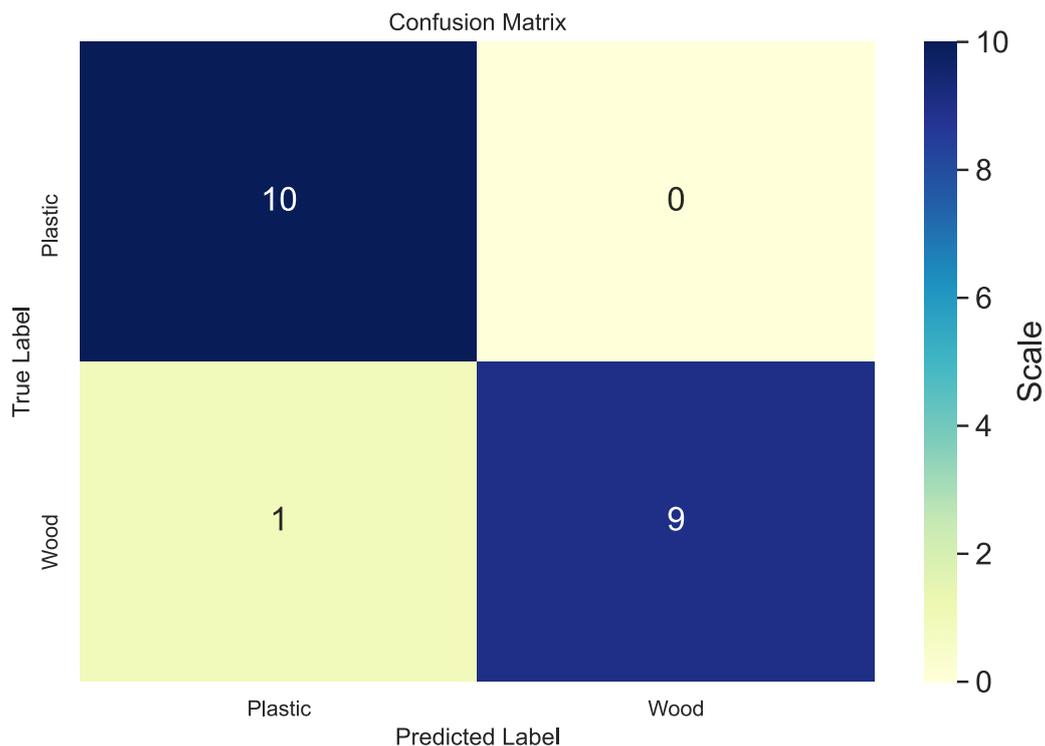


Figure 3: Confusion matrix using transfer learning CNN on PLAWO-40 dataset with pre-processing

compared to its use on PLAWO-40 dataset with the same learning strategy. We show an improvement of 0.58%, this is explained by the reduction of number of images the dataset (892 images in trashnet against 120 in PLAWO-40).

Similar to section 4.2.1, we show in Figure 5 and Figure 6 the confusion matrices for the two recognition strategies with pre-processing step. Good test accuracy score is achieved.

5 conclusion

In this paper, we proposed two learning strategies using Convolutional Neural Network (CNN) to recognize plastic waste texture and objects in the environment and especially in the deep ocean. We used two challenging datasets to test and improve the efficiency of our method. Our approach achieves high accuracy results and can be used with small challenging datasets. The obtained performance can be explained by the capability of CNN to work with small databases when using transfer learning strategies. We conclude that developing deep learning and computer vision agents can also help reduce pollution and manage waste. These agents can also provide a smart and green technology solution for waste in the wild.

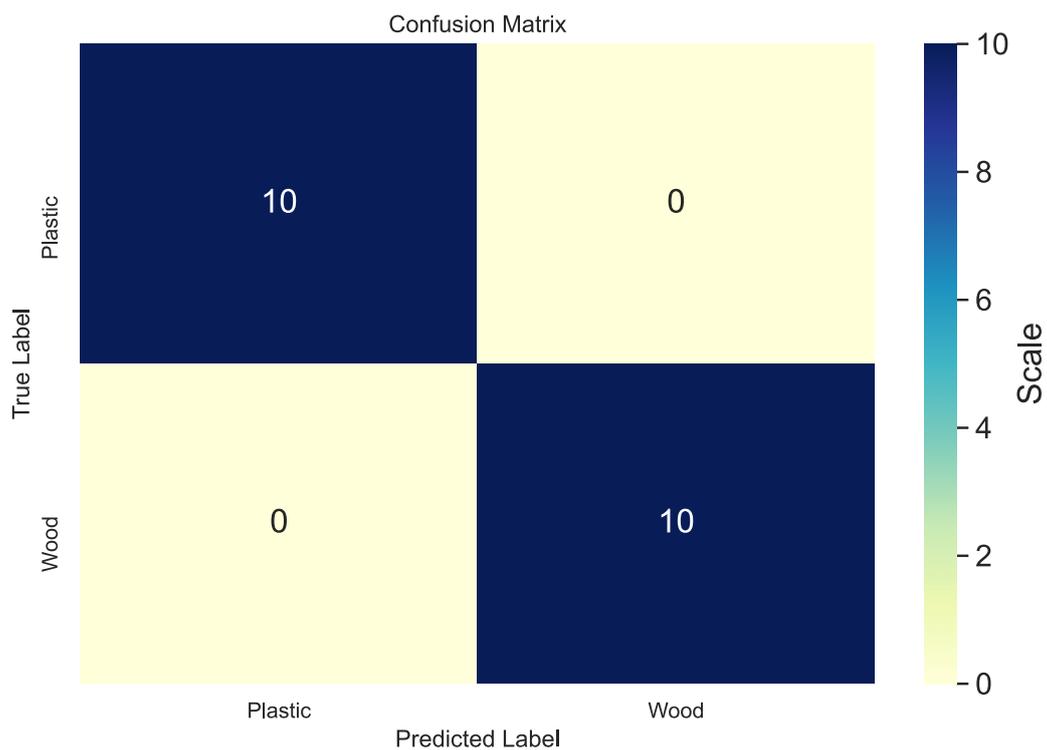


Figure 4: Confusion matrix using combined CNN-SVM on PLAWO-40 dataset with pre-processing

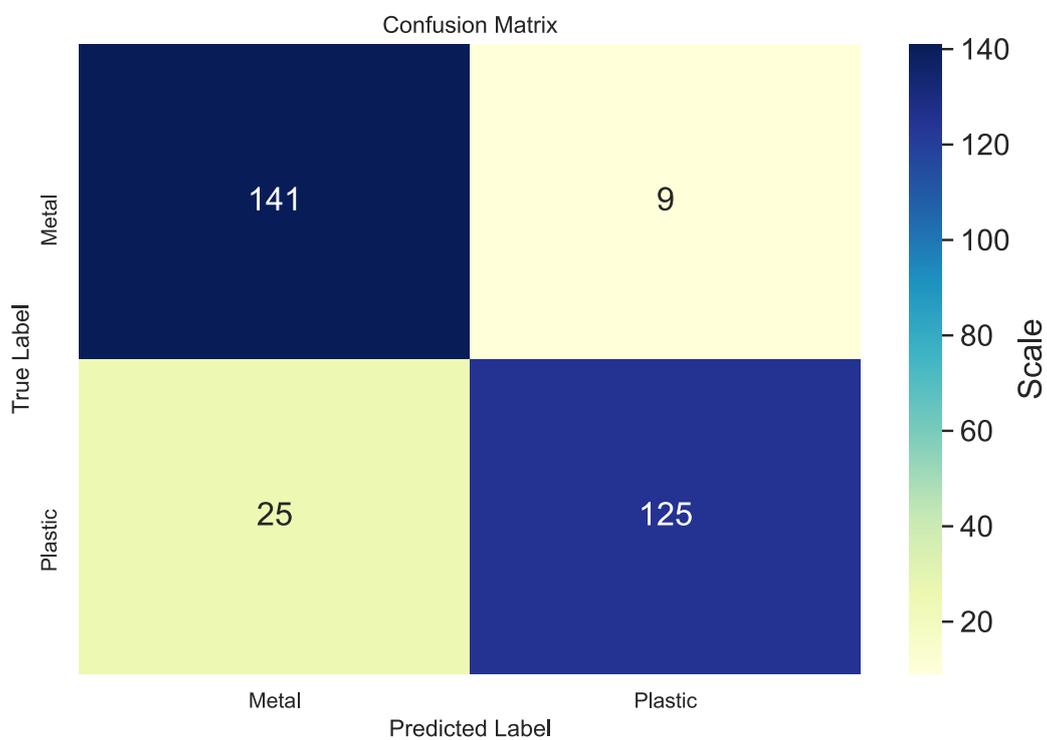


Figure 5: Confusion matrix using transfer learning CNN on trashnet dataset with pre-processing



Figure 6: Confusion matrix using combined CNN-SVM on trashnet dataset with pre-processing

References

- [1] O. Drzyzga, A. Prieto, *Microbial biotechnology* **12**, 66 (2019)
- [2] S. Dong, P. Wang, K. Abbas, *Computer Science Review* **40**, 100379 (2021)
- [3] C.C. Aggarwal et al., *Springer* **10**, 978 (2018)
- [4] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei, *Imagenet: A large-scale hierarchical image database*, in *2009 IEEE conference on computer vision and pattern recognition* (Ieee, 2009), pp. 248–255
- [5] A. Aoulalay, N. EL MAKHFI, M.C. ABOUNAIMA, M. MASSAR, *Classification of Moroccan Decorative Patterns Based on Machine Learning Algorithms*, in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)* (IEEE, 2020), pp. 1–7
- [6] C. Liu, L. Sharan, E.H. Adelson, R. Rosenholtz, *Exploring features in a bayesian framework for material recognition*, in *2010 IEEE computer society conference on computer vision and pattern recognition* (IEEE, 2010), pp. 239–246
- [7] I. Badami, K. Reinhard, *Material recognition: Bayesian inference or SVMs?*, in *Seminar on Computer Graphics (CESCG)* (Citeseer, 2012)
- [8] P. Bian, W. Li, Y. Jin, R. Zhi, *EURASIP Journal on Image and Video Processing* **2018**, 1 (2018)
- [9] P. Wieschollek, H. Lensch, *arXiv preprint arXiv:1609.06188* (2016)
- [10] F.A. Azis, H. Suhaimi, E. Abas, *Waste Classification using Convolutional Neural Network*, in *Proceedings of the 2020 2nd International Conference on Information Technology and Computer Communications* (2020), pp. 9–13
- [11] A. Krizhevsky, I. Sutskever, G.E. Hinton, *Communications of the ACM* **60**, 84 (2017)
- [12] M.D. Zeiler, R. Fergus, *Visualizing and understanding convolutional networks*, in *European conference on computer vision* (Springer, 2014), pp. 818–833
- [13] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, *Going deeper with convolutions*, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 1–9
- [14] K. Simonyan, A. Zisserman, *arXiv preprint arXiv:1409.1556* (2014)
- [15] K. He, X. Zhang, S. Ren, J. Sun, *Deep residual learning for image recognition*, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 770–778
- [16] S.J. Pan, Q. Yang, *22 (10): 1345* **1359** (2010)
- [17] L. Torrey, J. Shavlik, *IGI Global* **3**, 17 (2009)
- [18] M. Yang, G. Thung, *CS229 Project Report* **2016** (2016)