Using Class Activation Maps on Deep Neural Networks to Localise Waste Classifications

Created by Andrea Abela

Supervised by Mr. Thomas Gatt

Introduction

- Almost 40% of waste ends up in landfills instead of being recycled
- Waste mismanagement is an effect of our decreasing sense of duty
- The objective of this study is to alleviate the moral role of waste separation via waste detection
- This can be achieved through an automated system such as artificial intelligence (AI)

Research Questions

Is it possible for an AI system to be considered as accurate in classifying waste?

Is an AI system able to infer its results efficiently?

Is it possible to utilise specialised techniques such as weak supervision to integrate detection functionality?

Theory

Machine Learning (ML)

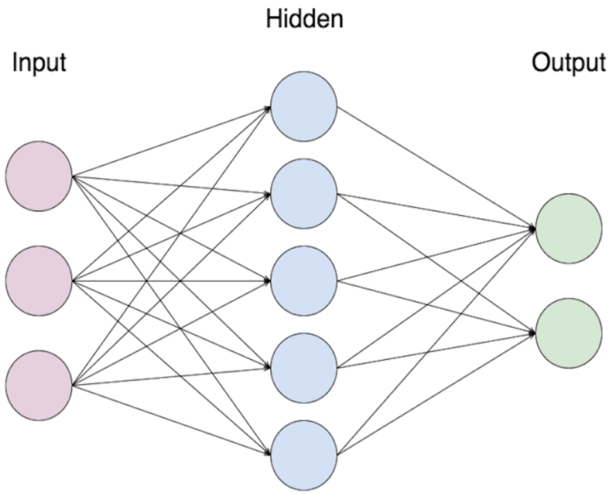
- Neural Network (NN)

Deep Learning (DL)

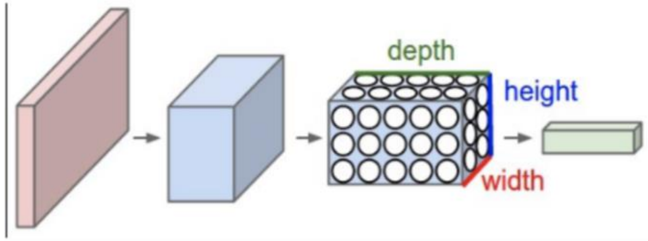
- Convolutional Neural Network (CNN)

Supervision

- Supervised Learning
- Weakly Supervised Learning



NN Diagram



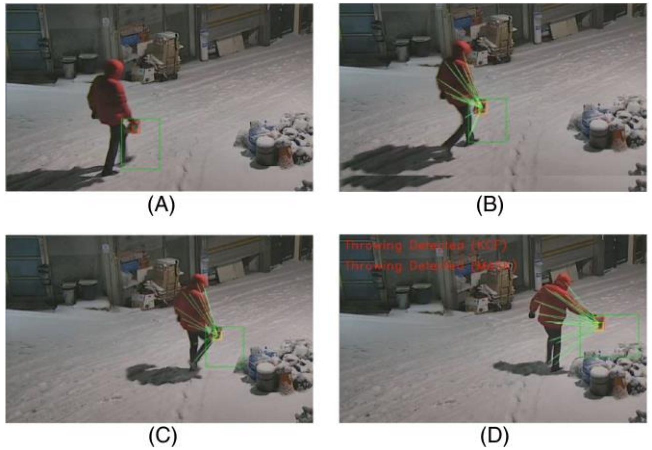
CNN Diagram

Related Works

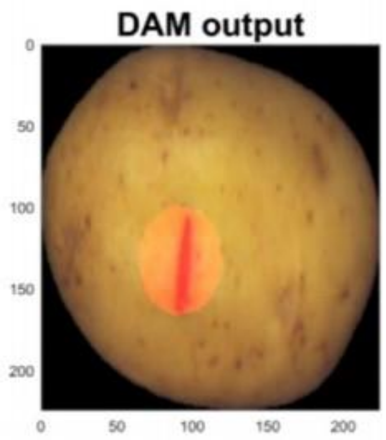
Studies that test deep neural network (DNN) models such as VGG, ResNet, MobileNet and DenseNet on waste classification use cases

Studies that attempt to create viable AI applications for waste management, such as illegal dumping detection and garbage pile marking

Other studies that test possibly useful techniques such as class activation maps (CAM) for weak supervision



Illegal Dumping Detection



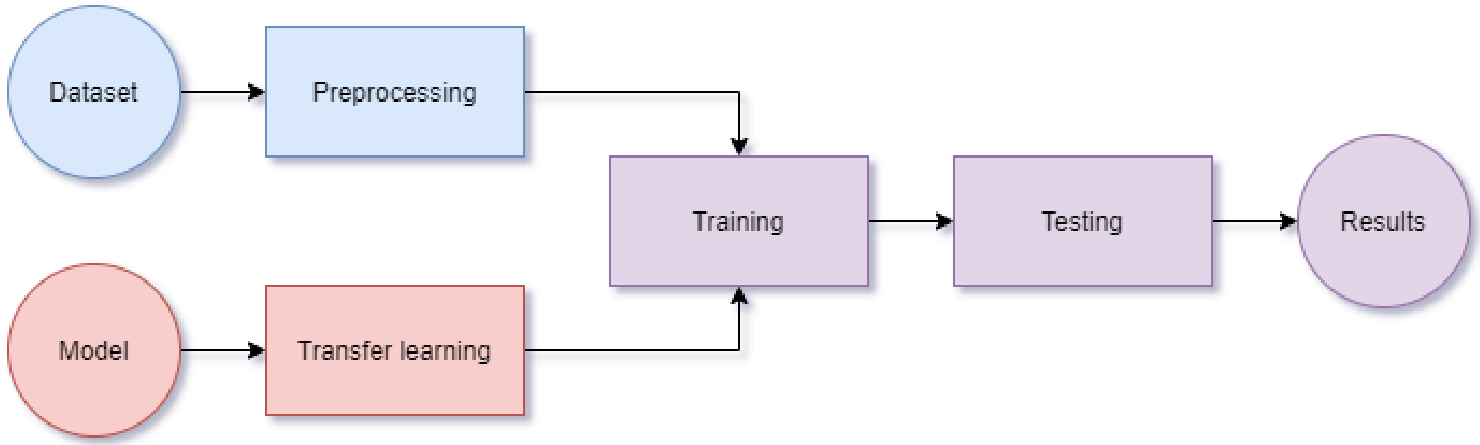
Potato Defect CAM

Pipeline

A pipeline inspired from previous studies was created

This was done to streamline the development process

It was also done to allow the existence of multiple configurations which give rise to multiple results



Datasets

TrashNet

Materials in Context (MINC)

Models

VGG16

ResNet50V2

MobileNetV2

DenseNet121

Fine-tuning

Classifier only

Low-level

High-level

Full re-training

Other techniques

CAMs

Image augmentation

Data splitting

Multiple Configurations

Dataset

Split into subsets
Pre-process images

Model

Load weights
Apply sensible fine-tuning

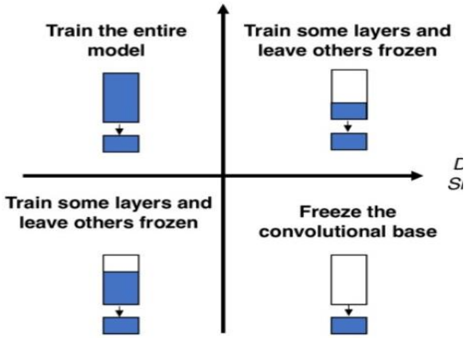
Train

Use stochastic gradient descent (SGD)
Multiple callbacks

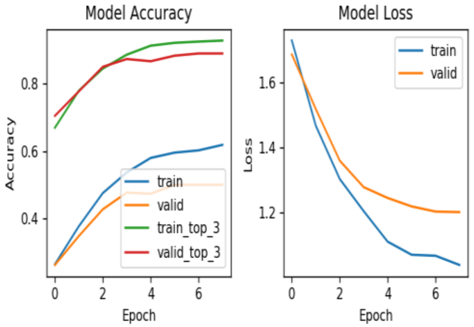
Test

Process predictions
Create bounding box CAMs
Generate metrics

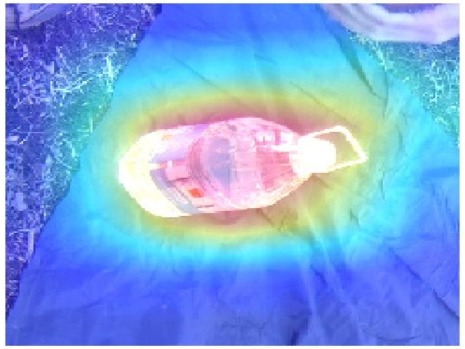
TrashNet sample



Size-similarity matrix



Training sample



MINC CAM sample

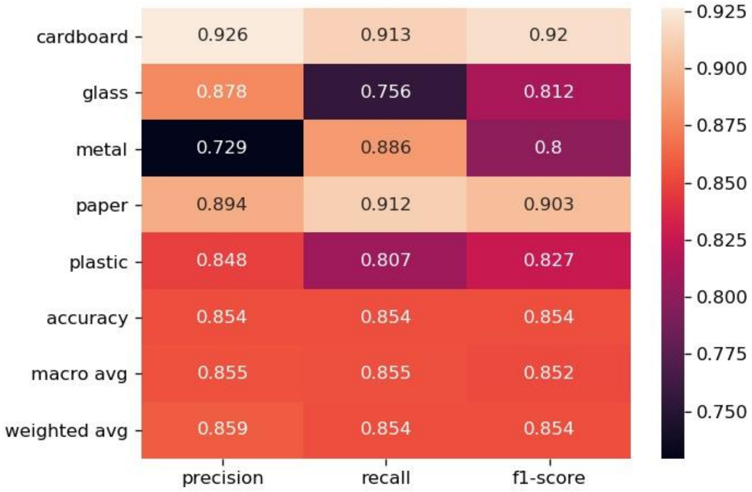
Proposed Method

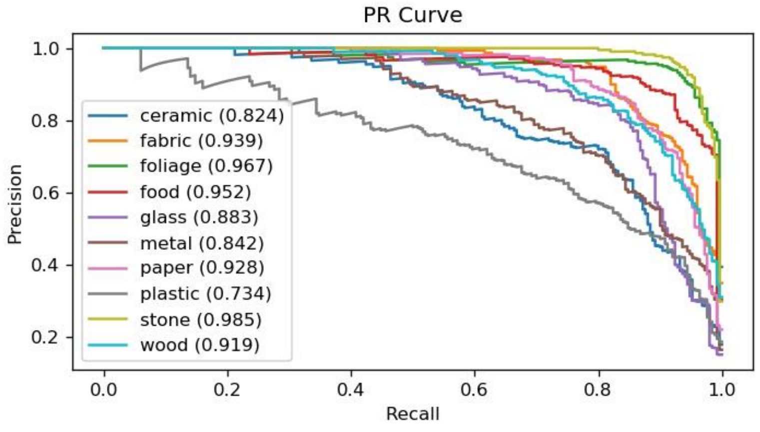
Accuracy Metrics

The metrics used were:

F1-score	PR-AUC score

Previous studies made use of accuracy, which is not predictive



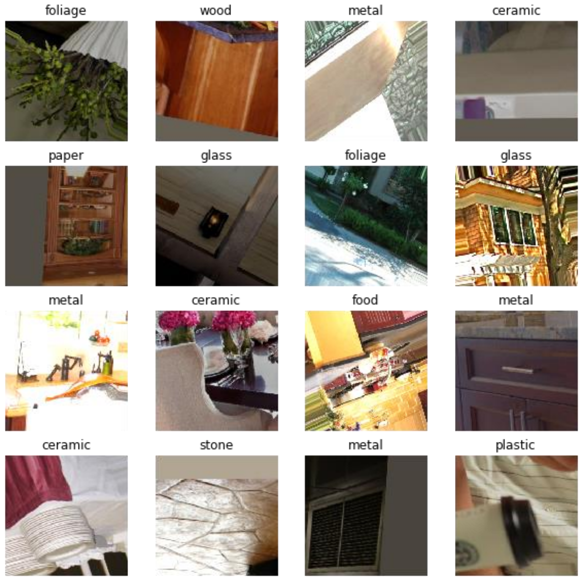


Average AUC over all classes: 0.897

Accuracy Results

TrashNet

The **low-level fine-tuned VGG16** model proved to be the most accurate with an **F1-score of 85.4%** and a **PR-AUC score of 93.3%**



MINC

The **fully re-trained DenseNet121** model gained the highest accuracy, by providing an **F1-score of 82.3%** and a **PR-AUC score of 89.7%**

Efficiency Metrics

The metrics used were:

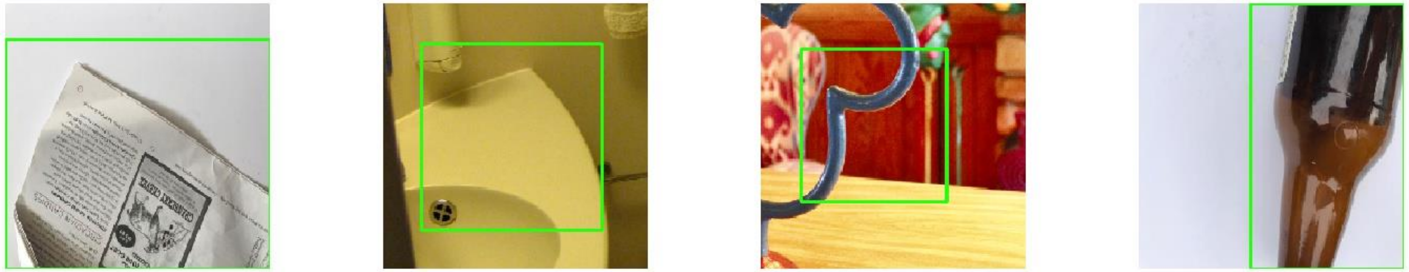
Average prediction and localisation time	Model weights file size

Not many previous studies used this method

Efficiency Results

- MobileNetV2 proved to be the most efficient algorithm
- It gained over 10 frames per second (FPS) in speed
- While maintaining under 10 megabytes (Mb) in size
- Despite not being the most accurate, it provided an F1-score of around 77% when fully re-trained on MINC
- If efficiency is a priority and a large dataset is available, MobileNetV2 is a solid choice

Localisation Metrics



Survey Bounding Box Samples

- No ground truth bounding boxes existed for localisation
- A survey was created to use a human benchmark as an evaluation method
- 15 images were chosen from the best models to represent each category from both datasets

Localisation Results

The low-level fine-tuned VGG16 trained on TrashNet obtained **83.2%** localisation accuracy

While the re-trained DenseNet121 MINC model gained **68.5%** localisation accuracy

Furthermore, it was found that the participants were confused with MINC, which is why it gained such a lower score

Considering that no ground truth was available, both models provided positive results

Results Review

It can be safely said that the objective was reached as accuracy, efficiency and localisation results proved to be positive

Improvements include applying more thorough fine-tuning methods and attempting techniques such as k-fold cross validation

Future work could also be undertaken by integrating video functionality and migrating the prototype to an embedded system

Research Paper

Since more work is planned on being done, a research paper was made

This was done as to prepare the study for SAMI 2021

This will provide a medium for the improvements and further work to be implemented

The End