



Design of Deep Learning-Based Models for Intelligent Waste Classification Using Single-Label and Multi-Label Approaches

Behnam Aminizadeh^{a*}, Zeinab Jahangiri^b

^aDepartment of Environmental Systems Engineering, Graduate Faculty of Environment, University of Tehran, Tehran, Iran.

*Corresponding Author: Behnam Aminizadeh behnamaminizadeh@ut.ac.ir, <https://orcid.org/0009-0009-5266-1306>

Article Info

Article type:
Research Article

Article history:
Received: 2026/01/28
Received in revised form:
2026/02/18
Accepted 2026/03/01
Available online: 2026/03/06

Keywords:
Transfer Learning,
Waste Classification,
Waste Management,
Intelligent Waste Sorting,
Deep Learning.

ABSTRACT

Today, with the rapid growth of technology, especially deep learning, the importance of using novel methods has become more evident than ever. Environmental issues such as waste management likewise require new, intelligent approaches that can reduce costs, accelerate operations, and minimize human error. Waste classification, as one of the fundamental and important aspects of waste management and recycling, is the main focus of this research. In this study, two deep learning models were trained on images from the Realwaste dataset, which includes about 4,200 images across the categories of metal waste, glass waste, paper waste, cardboard waste, food organic waste, plastic waste, vegetation waste, and textile waste, in order to advance intelligent waste sorting. The resulting models are a single-label model and a multi-label model. The base model uses the EfficientNetB0 architecture introduced by researchers at Google, but it was redesigned and trained through transfer learning to suit the specific problem addressed in this study. To increase data diversity, artificial data augmentation techniques such as rotation, lighting changes, and image blending using the CutMix and MixUp methods were employed. In addition, to address data imbalance, class weights were calculated based on the frequency of each class. The training and validation data were split at a ratio of 70% to 30%, and all programming was carried out in Python within the Visual Studio Code environment using the Scikit-learn and TensorFlow libraries. The experimental results showed that both models performed well in the intelligent classification of waste images. For images containing a single waste type, the single-label model sometimes provided higher confidence in identifying the correct category; however, the multi-label model, which exhibited better control of overfitting during training, delivered better overall performance. The multi-label model's average confidence in identifying the correct category in the experimental test, 0.875, was higher than that of the single-label model, 0.855. In the analysis of an image unrelated to waste, the multi-label model, with a confidence below 0.25, also made a smaller error than the single-label model, which had a confidence of 0.29. Finally, in the experimental test on images containing several waste types simultaneously, it was found that the EfficientNet architecture is not as reliable for analyzing images containing multiple waste types as it is for analyzing images containing a single waste type.

Cite this article: Aminizadeh, B., & Jahangiri, Z. (2026). Design of Deep Learning-Based Models for Intelligent Waste Classification Using Single-Label and Multi-Label Approaches. *Strategies of Engineering Sciences and Sustainable Systems Management*, 1(1), 22-40. <https://doi.org/10.22034/jsesm.2026.252591>



Publisher: Iranian Scientific Association of Waste Engineering and Management.

DOI: <https://doi.org/10.22034/jsesm.2026.252591>

Introduction

Rapid progress in artificial intelligence, machine learning, and especially deep learning has transformed how unstructured data, particularly images, are analyzed, with convolutional neural networks remaining one of the main foundations of image classification. Environmental challenges such as excessive waste generation call for intelligent solutions, and waste management and sorting have drawn considerable global attention, as population growth, urbanization, and industrial expansion increase household, commercial, industrial, and agricultural waste volumes. Traditional waste separation is usually manual and rule-based, an approach difficult to scale and labor-intensive; applying artificial intelligence, by contrast, can increase classification accuracy and reduce operating cost and time, with approaches such as smart bins and sorting robots proposed as effective directions. Building on this motivation, the present study develops a single-label model and a multi-label model, both based on EfficientNetB0 and transfer learning, trained on RealWaste data, and compares their performance across a range of test images. EfficientNet was chosen as the base architecture because this family of models balances depth, width, and image resolution well, and EfficientNetB0, despite far fewer parameters than earlier architectures such as VGG16, is generally a more optimized and accurate model. Persian-language and domestic research specifically centered on intelligent waste sorting with the help of deep learning and image processing is not abundant, and the present study also aims to inform future research in this direction, including the potential contribution of object-detection models such as YOLO for images containing several waste types at once.

Method

The RealWaste dataset, containing more than 4,700 images across nine categories, was released in late 2023 and used as the basis for this study. The miscellaneous-trash category was excluded so training could focus on precisely labeled images, leaving eight target categories: metal, glass, paper, cardboard, food organics, plastic, vegetation, and textile waste, comprising roughly 4,200 images. Further images matching each category were downloaded from the internet, bringing the working image pool to approximately 5,000 images. Because RealWaste images generally each focus on a single waste item, single-label images were also combined into composite images to give the multi-label model better training conditions, alongside rotation and brightness variation; class weights based on category frequency were applied, since image counts differ considerably across categories. For both models, EfficientNetB0, pretrained on roughly 1.2 million images across about 1,000 classes, served as the starting point, so it's already-learned visual features could be reused here; the original 1,000-class output layer was replaced with an 8-class layer. Early layers, capturing general image features, were kept frozen while newly added layers trained on the project's data; some previously frozen layers were later unfrozen for closer adaptation to waste imagery. Both models used independent sigmoid outputs, so each class receives a value between 0 and 1 that need not sum to 1; this same design was used for the single-label model too, keeping the two architectures consistent for easier comparison. Accordingly, the single-label model reports only the label with the highest value, while the multi-label model displays every label exceeding a 0.25 threshold, adjustable in the model's test code. Training and validation data were split 70:30; Adam was used instead of RMSProp, originally used for EfficientNetB0, for more stable, faster convergence on this smaller dataset; all programming was done in Python, in Visual Studio Code, using Scikit-learn and TensorFlow. Over training, the single-label model showed some widening between training and validation loss, a mild but not severe degree of overfitting; the multi-label model, trained with fewer epochs partly in response, showed a more balanced loss, without clear overfitting or underfitting.

Results

Both trained models were tested on new images not seen during training, downloaded from the internet. The first four test images each showed a single type of waste — paper, textile waste, food organics, and plastic — and both models identified the correct category in every case, though confidence values differed. On the paper image, the single-label model reported higher confidence (0.84 versus 0.76); on the remaining three, the multi-label model reported higher confidence (0.87 versus 0.80 for textile waste, 0.93 versus 0.87 for food organics, and 0.94 versus 0.91 for plastic). On the food organics image, the multi-label model produced two labels, food organics and textile waste, with the correct category, food organics, receiving much higher confidence (0.93) than the second label (0.31). Averaged across these four images, the multi-label model's confidence in the correct category was 0.875, compared with 0.855 for the single-label model. On an image unrelated to any of the eight trained categories, the multi-label model's confidence stayed below 0.25 for every category, while the single-label model reported 0.29 confidence for one category, metal; the multi-label model therefore performed better on this image. On a composite image formed by combining five different waste types, the multi-label model correctly identified two categories, food organics and glass, at 0.35 and 0.30 confidence, respectively; one of these, food organics, was also identified by the single-label model, at 0.28 confidence, and the multi-label model's higher confidence on this shared category is why it performed better on this image. Overall, confidence levels were low on this image containing several waste types at once, a pattern suggesting that object-detection architectures such as YOLO may be better suited to such images.

Conclusions

Both models performed adequately in classifying waste images, and the multi-label model was overall more accurate than the single-label model across the test images used. The multi-label model also performed better on the image unrelated to the trained categories and on the composite image containing several waste types, although the single-label model occasionally reported higher confidence on individual single-waste-type images. Because of its greater flexibility in identifying more than one class in a single image, the multi-label approach appears better suited to real-world recycling lines that handle a wide variety of waste, though the EfficientNetB0 architecture, in either its single-label or multi-label form, is not a fully reliable choice for analyzing complex images containing multiple waste types at once. Beyond the models themselves, this comparison of single-label and multi-label performance on a waste-classification task adds to the literature on deep learning applications in environmental management, and, subject to further validation, the developed models could support recycling lines and smart waste-management systems, helping reduce cost, speed up operations, and reduce human error. Future work in this direction could focus on increasing the size and diversity of the training data, particularly in lighting, angle, and image quality; using genuinely composite images, rather than synthetically combined ones, to further improve the multi-label model for real-world scenarios; incorporating object-detection architectures such as YOLO for complex, multi-waste images; and, before training, setting aside a separate test set on which to report standard metrics such as accuracy, precision, recall, and F1, allowing a more rigorous evaluation than the qualitative testing carried out here.



طراحی مدل های مبتنی بر یادگیری عمیق برای تفکیک هوشمند پسماند با رویکردهای تک برچسبی و چندبرچسبی

بهنام امینی زاده^{a*}، زینب جهانگیری^a

^aگروه مهندسی سیستم های محیط زیست، دانشکده تحصیلات تکمیلی محیط زیست، دانشگاه تهران، تهران، ایران.

*نویسنده مسئول: بهنام امینی زاده behnamaminizadeh@ut.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	
تاریخ دریافت: ۱۴۰۴/۱۱/۰۸	
تاریخ بازنگری: ۱۴۰۴/۱۱/۲۹	
تاریخ پذیرش: ۱۴۰۴/۱۲/۱۰	
تاریخ انتشار: ۱۴۰۴/۱۲/۱۵	
کلیدواژه ها: انتقال یادگیری، تفکیک زباله، مدیریت پسماند، هوشمندسازی، یادگیری عمیق.	امروزه با رشد سریع تکنولوژی، به خصوص یادگیری عمیق، اهمیت به کارگیری روش های نوین بیش از پیش آشکار شده است. مسائل محیط زیستی همچون مدیریت پسماند نیز نیازمند رویکردهای جدید و هوشمند است تا باعث کاهش هزینه ها، تسریع در انجام امور و کاهش خطای انسانی شود. طبقه بندی پسماند، به عنوان یکی از جنبه های اساسی و مهم در مدیریت و بازیافت پسماند، محور اصلی این پژوهش است. در این پژوهش با آموزش دو مدل یادگیری عمیق بر روی تصاویر مجموعه ی Realwaste که حدود ۴۲۰۰ تصویر از دسته های پسماندهای فلزی، پسماندهای شیشه ای، پسماندهای کاغذی، پسماندهای مقوایی، پسماندهای آلی غذایی، پسماندهای پلاستیکی، پسماندهای گیاهی و پسماندهای نساجی/پارچه ای هستند، در مسیر تفکیک هوشمند پسماند گام برداشته شده است. مدل های حاصل از این پژوهش تک برچسبی و چندبرچسبی هستند. مدل پایه از معماری EfficientNetB0 بهره گرفته، اما با روش انتقال یادگیری، متناسب با مسئله ی خاص این پژوهش، بازطراحی و آموزش داده شده است. به منظور افزایش تنوع داده ها، تکنیک های افزایش مصنوعی داده ها مانند چرخش، تغییر نور و ترکیب تصاویر با روش های CutMix و MixUp به کار گرفته شد. همچنین برای مقابله با عدم توازن داده ها، وزن دهی کلاس ها بر اساس فراوانی هر کلاس محاسبه گردید. داده های آموزش و اعتبارسنجی با نسبت ۷۰ به ۳۰ درصد تقسیم شدند و برنامه نویسی به زبان پایتون در محیط Visual Studio Code با استفاده از کتابخانه های Scikit-learn و TensorFlow انجام گرفت. نتایج آزمون های تجربی نشان دادند هر دو مدل در تفکیک هوشمند تصاویر پسماند عملکرد خوبی دارند. مدل تک برچسبی در تصاویر با یک نوع پسماند گاهی اطمینان بالاتری در تشخیص دسته ی صحیح ارائه داد اما مدل چندبرچسبی که بیش برآزش کنترل شده تری در حین آموزش داشت، در مجموع عملکرد بهتری ارائه کرد؛ متوسط اطمینان مدل چندبرچسبی در تشخیص دسته ی صحیح در آزمون تجربی برابر با ۰/۸۷۵ نیز از متوسط اطمینان مدل تک برچسبی در تشخیص دسته ی صحیح برابر با ۰/۸۵۵ بیشتر بود. در تحلیل تصویر نامرتبط به پسماند نیز مدل چندبرچسبی با اطمینان کمتر از ۰/۲۵ از مدل تک برچسبی با اطمینان ۰/۲۹ خطای کمتری داشت. در آزمون تجربی تحلیل تصاویر با حضور هم زمان چند نوع پسماند، مشخص شد معماری EfficientNet، به اندازه ی تحلیل تصاویر با حضور یک نوع پسماند، قابل اعتماد نیست.

استناد به این مقاله: امینی زاده، بهنام و جهانگیری، زینب. (۱۴۰۴). طراحی مدل های مبتنی بر یادگیری عمیق برای تفکیک هوشمند پسماند با رویکردهای تک برچسبی و چندبرچسبی. راهبردهای علوم مهندسی و مدیریت سیستم های پایدار، ۱(۱)، ۴۰-۲۲. <https://doi.org/10.22034/jscsm.2026.252591>



© نویسندگان

DOI: <https://doi.org/10.22034/jscsm.2026.252591>.

ناشر: انجمن علمی مهندسی و مدیریت پسماند ایران

مقدمه

در دهه‌های اخیر، پیشرفت شتابان فناوری‌های نوظهور به ویژه در حوزه‌های هوش مصنوعی، یادگیری ماشین و یادگیری عمیق، شیوه تحلیل داده‌های غیرساختاریافته، به خصوص تصاویر را به طور چشمگیری متحول کرده است. مرورها نشان می‌دهند که شبکه‌های عصبی کانولوشنی همچنان یکی از پایه‌های اصلی طبقه‌بندی تصویر هستند و یادگیری عمیق نیز به عنوان یک چارچوب فراگیر برای حل مسائل واقعی در حوزه‌های گوناگون به کار می‌رود (Chen et al., 2021; Sarker, 2021).

در سال‌های اخیر، محققان داخلی استفاده از الگوریتم‌های یادگیری ماشین در حل مسائل محیط‌زیستی را در دستور کار خود قرار داده‌اند (Mohammadzadeh et al., 2026; Aghdam et al., 2024). در زمینه‌های مختلف و پژوهش‌های مختلف از الگوریتم‌های یادگیری ماشین به شکل موثر و کارآمد استفاده می‌شود. این الگوریتم‌ها در مسائل مختلف محیط‌زیستی نیز نقش مهمی ایفا می‌کنند. به عنوان مثال در پیش‌بینی کیفیت خروجی فاضلاب تصفیه‌خانه‌ی شهری از یادگیری ماشین بهره برده شده است (Biglarijoo et al., 2025). استفاده از پردازش تصویر و یادگیری عمیق نیز یکی از حوزه‌هایی است که می‌تواند بیشتر مورد توجه محققان داخلی جهت حل مشکلات محیط زیستی قرار بگیرد چراکه استفاده از پردازش تصویر در دنیای امروز، بسیار مورد توجه محققان بین‌المللی قرار گرفته و نتایج خوبی نیز به دنبال داشته است (Zhu et al., 2017; Ma et al., 2019; Yuan et al., 2020; Himeur et al., 2022; Bai et al., 2023). یادگیری ماشین به عنوان یکی از زیرشاخه‌های هوش مصنوعی، به سیستم‌ها امکان می‌دهد بدون برنامه‌نویسی صریح و با تکیه بر داده‌ها و نمونه‌ها، الگوها را بیاموزند و عملکرد خود را بهبود بخشند؛ با توسعه یادگیری عمیق، به ویژه شبکه‌های عصبی چندلایه، توان مدل‌های هوش مصنوعی در تحلیل داده‌های تصویری، متن و ویدیو به طور چشمگیری افزایش یافته است (De Mauro, 2022).

یکی از حوزه‌هایی که می‌تواند از قابلیت‌های هوش مصنوعی سود بسیاری ببرد، مسائل مرتبط با محیط زیست و توسعه‌ی پایدار است. چالش‌های محیط زیستی نظیر تغییرات اقلیمی، کاهش منابع طبیعی، آلودگی‌های محیط زیستی و تولید بیش از حد پسماند، نیازمند راهکارهایی نوآورانه و دقیق هستند (Alotaibi & Nassif, 2024; Dhiman et al., 2020; Vinuesa et al., 2024). از جمله مسائل محیط زیستی مهم که توجه جهانی را به خود جلب کرده، مسئله‌ی مدیریت پسماند و تفکیک زباله است. افزایش جمعیت، توسعه‌ی شهری، مصرف‌گرایی و گسترش صنایع، منجر به تولید حجم انبوهی از زباله‌های خانگی، تجاری، صنعتی و کشاورزی شده است (Kaza et al., 2018; Chen, 2018). در روش‌های سنتی مدیریت پسماند، جداسازی و دسته‌بندی معمولاً به صورت دستی و با قواعد ثابت انجام می‌شود؛ این رویکردها در برابر حجم بالای زباله، تنوع زیاد مواد و نیاز به دقت بالا، عموماً کم‌مقیاس، نیازمند نیروی انسانی قابل توجه و کم‌انعطاف هستند (Dawar et al., 2025).

در مقابل، به کارگیری هوش مصنوعی در مدیریت پسماند می‌تواند دقت طبقه‌بندی را افزایش دهد، هزینه و زمان عملیات را کاهش دهد و بازیابی منابع را بهبود بخشد؛ در مرورهای اخیر، کاربردهایی مانند سطل‌های هوشمند، ربات‌های تفکیک زباله، پایش مبتنی بر حسگر و مدل‌های پیش‌بینی تولید پسماند به عنوان مسیرهای موثر مطرح شده‌اند (Fang et al., 2023). در مدیریت پسماند، خودکارسازی تفکیک زباله به یک مسئله‌ی کاربردی و مهم تبدیل شده است، زیرا حجم پسماند شهری رو به افزایش است، تفکیک دستی پرهزینه و زمان‌بر است و آلودگی جریان بازیافت می‌تواند کارایی سامانه‌های بازیافت را کاهش دهد. یک مرور نظام‌مند جدید نشان می‌دهد که کاربردهای یادگیری ماشین و یادگیری عمیق در مدیریت پسماند شهری به سرعت در حال گسترش است (Dawar et al., 2025). بنابراین، ترکیب یادگیری عمیق با سامانه‌های تفکیک پسماند می‌تواند هم از نظر عملیاتی و هم از نظر اقتصادی، جایگزینی دقیق‌تر و

مقیاس‌پذیرتر برای روش‌های سنتی باشد (Ahmad et al., 2025; Sayem et al., 2025). برای ارزیابی الگوریتم‌ها در شرایط نزدیک به واقعیت، استفاده از مجموعه داده‌های واقعی اهمیت ویژه‌ای دارد، Realwaste یکی از مجموعه داده‌های واقعی و نسبتاً جدید پسماند است و امکان آزمون مدل‌ها را بر روی نمونه‌هایی فراهم می‌کند که به کاربرد عملی نزدیک‌اند (Single et al., 2023).

از سوی دیگر، برای انتخاب معماری پایه، EfficientNet یک گزینه‌ی مناسب به شمار می‌آید؛ زیرا این خانواده از مدل‌ها تعادل مناسبی میان عمق، عرض و وضوح تصویر ایجاد می‌کند و در مسائل تصویری کارآیی بالایی دارد. در نتیجه، EfficientNetB0 می‌تواند برای طبقه‌بندی پسماند در مسئله‌ی حاضر به کار رود (Tan & Le, 2019).

در کنار طبقه‌بندی، برای تصاویر شلوغ و چندشیئی نیز روش‌های تشخیص شیء اهمیت دارند؛ مطالعات جدید نشان می‌دهند که مدل‌هایی مانند YOLOv8 می‌توانند برای شناسایی و تفکیک پسماند در سناریوهای واقعی مفید باشند و در کنار مدل‌های طبقه‌بندی‌محور، یک مسیر مکمل برای تحلیل تصاویر پیچیده فراهم کنند (Arishi, 2025; Mehadjbia & Slaoui-Hasnaoui, 2024).

انتقال یادگیری رویکردی است که در آن دانش یا الگوهای آموخته شده از یک دامنه‌ی مبدا به یک دامنه‌ی هدف منتقل می‌شود تا مدل جدید با داده‌ی کمتر، سریع‌تر و کارآمدتر آموزش ببیند (Wu & He, 2025).

در آموزش مدل‌های بینایی ماشین، افزایش مصنوعی داده‌ها یکی از راهبردهای استاندارد برای افزایش تنوع نمونه‌های آموزشی و کاهش بیش‌برازش است؛ روش‌هایی مانند چرخش، برش، تغییر رنگ و ترکیب تصاویر در کاربردهای طبقه‌بندی تصویر موثر گزارش شده‌اند. در مسائلی مانند طبقه‌بندی پسماند، کمبود داده‌های برچسب خورده و ناهمگونی نمونه‌ها معمولاً باعث می‌شود مدل‌های عمیق برای تعمیم‌پذیری به راهبردهای مکمل نیاز داشته باشند. در چنین شرایطی، انتقال یادگیری با انتقال دانش از یک مدل از پیش آموزش دیده به دامنه‌ی هدف، وابستگی به داده‌های جدید را کاهش می‌دهد؛ از سوی دیگر، افزایش مصنوعی داده‌ها در تصاویر، با بالا بردن تنوع نمونه‌های آموزشی، به کاهش بیش‌برازش و بهبود عملکرد مدل کمک می‌کند. بنابراین، ترکیب انتقال یادگیری و افزایش مصنوعی داده‌ها، در کاربردهای بینایی ماشین، به ویژه در طبقه‌بندی تصاویر واقعی پسماند، یک راهبرد موثر برای رسیدن به دقت بالاتر و آموزش کارآمدتر است (Zhuang et al., 2020; Jordan & Mitchell, 2015).

بررسی‌ها نشان می‌دهد پژوهش‌های داخلی و مقالات فارسی منتشر شده با محوریت مسئله‌ی تفکیک هوشمند پسماند با کمک یادگیری عمیق و پردازش تصویر زیاد نیستند؛ در یکی از این پژوهش‌ها، محققان با کمک بیش از ۳۵۰۰ تصویر جمع‌آوری شده از اینترنت برای چهار نوع از بطری‌ها و با استفاده از معماری پایه VGG16 مدلی برای طبقه‌بندی تصاویر طراحی کردند (Bayatzadeh et al. 2025). در پژوهش حاضر ضمن استفاده از مجموعه داده‌ای با تعداد تصاویر و تنوع تصاویر بیشتر، از معماری مدرن‌تر EfficientNetB0 بهره برده شده است. پژوهش‌های قبلی نشان از این دارند که EfficientNetB0 باوجود تعداد پارامترهای خیلی کم‌تر نسبت به VGG16، مدلی بهینه‌تر و معمولاً دقیق‌تر است (Ali et al, 2025; Tan & Le, 2019). پژوهش حاضر در تلاش است با ارائه‌ی مدل‌ها و راهکارهایی برای ارتقاء کیفیت مدل‌ها جهت هوشمندسازی طبقه‌بندی پسماند، اهدافی مانند کاهش هزینه‌ها، تسریع در انجام امور و کاهش خطای انسانی را محقق کند.

روش‌شناسی پژوهش

مجموعه داده Realwaste که دارای بیش از ۴۷۰۰ تصویر از ۹ دسته مختلف است، در اواخر سال ۲۰۲۳ منتشر شده است. جهت انجام پژوهش، دسته‌ی متفرقه حذف شد تا تمرکز مدل بر روی تصاویر با برچسب دقیق حفظ شود

اما این قابلیت موجود است که اگر مدل در تشخیص برچسب تصویری، از اطمینان کافی برخوردار نبوده، تصویر مورد نظر در دسته‌ی متفرقه قرار گیرد. بنابراین دسته‌های باقی مانده عبارت‌اند از پسماندهای فلزی، پسماندهای شیشه‌ای، پسماندهای کاغذی، پسماندهای مقوایی، پسماندهای آلی غذایی، پسماندهای پلاستیکی، پسماندهای گیاهی و پسماندهای نساجی/پارچه‌ای؛ این داده‌ها پس از حذف داده‌های دسته متفرقه حدود ۴۲۰۰ عدد شدند (Single et al., 2023). جدول ۱، اطلاعات مجموعه داده را نمایش می‌دهد.

جدول ۱. اطلاعات مجموعه داده Realwaste

Table 1. Realwaste dataset information

ردیف	عنوان اصلی موجود در مجموعه داده	ترجمه فارسی	تعداد تصاویر موجود در مجموعه داده
۱	Cardboard	پسماندهای مقوایی	۴۶۱
۲	Food Organics	پسماندهای آلی غذایی	۴۱۱
۳	Glass	پسماندهای شیشه‌ای	۴۲۰
۴	Metal	پسماندهای فلزی	۷۹۰
۵	Paper	پسماندهای کاغذی	۵۰۰
۶	Plastic	پسماندهای پلاستیکی	۹۲۱
۷	Textile Trash	پسماندهای نساجی/پارچه‌ای	۳۱۸
۸	Vegetation	پسماندهای گیاهی	۴۳۶
۹	Miscellaneous Trash	پسماندهای متفرقه	۴۹۵

شکل ۱ دربرگیرنده تعدادی از نمونه تصاویر هر دسته از مجموعه Realwaste می‌باشد.



شکل ۱. نمونه‌ای از تصاویر در مجموعه داده Realwaste

Figure 1. Example of images in the Realwaste dataset

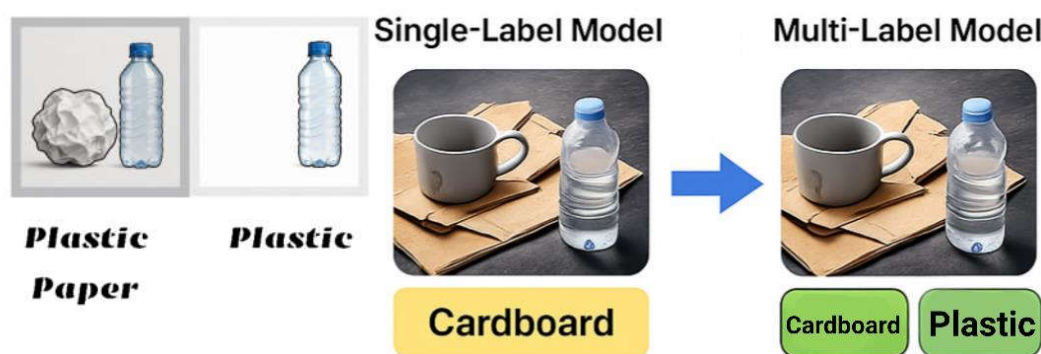
در شکل ۱، شماره هر دسته با ردیف‌های جدول ۱ متناظر می‌باشد. در هر دسته، تعدادی تصویر جدید (جدا از داده‌های اولیه) مطابق با آن دسته، برای کمک به افزایش بیشتر تنوع داده‌ها از اینترنت دانلود و به دسته‌ها اضافه شد.

تصاویر اولیه به حدود ۵۰۰۰ تصویر رسیدند. سپس تکنیک‌های افزایش مصنوعی داده‌ها^۱ مانند چرخش، تغییر نور، ترکیب تصاویر در نظر گرفته شد تا سعی شود از بیش‌برازش^۲ جلوگیری شود. این تکنیک‌های افزایش مصنوعی داده‌ها در کد پایتون آموزش مدل، کدنویسی شده قرار گرفتند.

در این پژوهش، دو مدل ساخته شده است؛ یک مدل تک‌برچسبی که برای هر تصویر، یک دسته را مشخص می‌کند (یکی از ۸ دسته‌ی اصلی) و مدل دوم، مدل چندبرچسبی است که می‌تواند برای هر تصویر و برای تمام دسته‌ها (هر ۸ دسته) مستقل از سایر دسته‌ها، بررسی کند و برای دسته‌هایی که مدل از اطمینان قابل قبولی برخوردار است، برچسب را اطلاق کند. با این روش می‌توان تلاش کرد اگر در یک تصویر، به عنوان مثال کارتن موجود بود و پلاستیک هم موجود بود، مدل هر دو دسته را مشخص کند و نمایش دهد:

پسماندهای مقوایی - پسماندهای پلاستیکی

شکل ۲، نمونه‌هایی از تفاوت عملکرد دو مدل را نمایش می‌دهند.



شکل ۲. تفاوت مدل تک‌برچسبی و چندبرچسبی

Figure 2. Difference between single-label and multi-label models

در شکل ۲ و تصویر سمت راست، مدل تک‌برچسبی آن دسته‌ای را خروجی می‌دهد که اطمینانش روی آن بیشتر است که در اینجا Cardboard انتخاب شده است؛ اما انتظار می‌رود مدل چندبرچسبی هر دو برچسب صحیح را خروجی دهد. در تصویر سمت چپ در شکل ۲، سمت راست فقط یک تصویر قرار دارد، بنابراین انتظار می‌رود هر دو مدل یک برچسب خروجی دهند. اما تصویر سمت چپ، دو نوع تصویر دارد بنابراین انتظار می‌رود مدلی که توانایی خروجی دادن بیش از یک برچسب را دارد، هر دو برچسب صحیح را خروجی دهد. لازم به ذکر است که تصاویر موجود در شکل ۲، ارتباطی با مجموعه داده‌ی Realwaste ندارند و صرفاً برای شفافیت بهتر در تفاوت رویکرد دو مدل آورده شده‌اند.

از آنجا که تصاویر مجموعه Realwaste روی یک زباله تمرکز دارند، سعی شده است با تکنیک‌های ترکیب تصاویر، شرایط بهتری برای آموزش مدل چندبرچسبی ایجاد شود. در این پروژه از مدل EfficientNetB0 برای نقطه شروع آموزش، کمک گرفته شده، اما تغییرات لازم و مطابقت دادن با داده‌های خاص این پروژه، اعمال شده است؛ در ادامه توضیحاتی ارائه می‌گردد. این مدل معرفی شده توسط پژوهشگران شرکت گوگل، روی حدود ۱/۲ میلیون تصویر با حدود ۱۰۰۰ دسته، آموزش دیده است. اما در پاسخ به این سوال که این مدل چه کمکی می‌تواند به این پروژه‌ی خاص با ۸ دسته، جهت تفکیک زباله کند؟ می‌توان گفت: از این مدل استفاده شده تا از ویژگی‌های عمومی و پایه‌ای تصاویر

¹ Data Augmentation

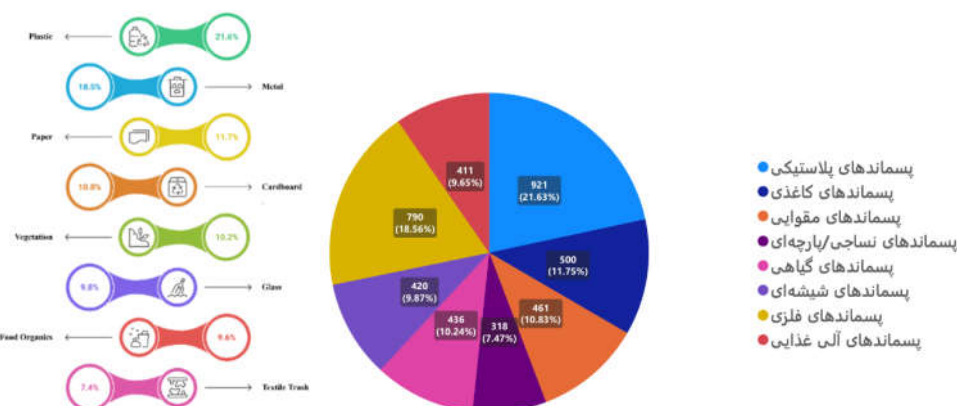
² Overfitting

که مدل روی داده‌ها با تعداد زیاد آموخته است (مثل تشخیص لبه‌ها، اشکال، الگوهای پایه‌ای) بهره برده شود. اما مدل جدید، برای یک مسئله‌ی خاص طراحی شده و روی حدود ۵۰۰۰ تصویر آماده شده آموزش دیده است (مجموعه داده‌ی Realwaste و تصاویر اضافه شده به آن). مدل چندبرچسبی آماده شده با عملکرد مدل پژوهشگران شرکت گوگل که تک‌برچسبی است، متفاوت است.

یک شبکه عصبی (مثل EfficientNetB0) از تعداد زیادی لایه تشکیل شده است که هر لایه، وظیفه‌ای دارد (Khanam et al., 2024). مدل معرفی شده پژوهشگران گوگل، بیش از ۲۰۰ لایه دارد. از این مدل کمک گرفته شده، اما نه برای طبقه‌بندی در ۱۰۰۰ دسته بلکه برای طبقه‌بندی در ۸ دسته‌ی مسئله‌ی خاص مورد بررسی در این پژوهش، کمک گرفته شده است. به نوعی از این مدل به عنوان یک نقطه شروع استفاده شده است، اما تغییراتی اعمال شده تا برای مسئله‌ی خاص مورد بررسی، مناسب گردد. لایه‌هایی حذف شده اند و لایه‌هایی اضافه شده اند و مدل به جای ۱۰۰۰ دسته، ۸ دسته تشخیص می‌دهد. انجام این تغییرات با کدنویسی به زبان پایتون صورت گرفته است. داده‌ها با تکنیک‌هایی مثل CutMix و MixUp متنوع‌تر شده اند.

مدل EfficientNet، می‌تواند در یاد گرفتن ویژگی‌های عمومی (که برای تشخیص تمام و یا اکثر تصاویر کاربردی است) مثل لبه‌ها مفید باشد. ابتدا برخی لایه‌ها، غیرقابل تغییر و ثابت شدند و سپس لایه‌های اضافه شده، آموزش دیدند (روی داده‌های این پروژه) و سپس برخی لایه‌های ثابت شده و فریز شده آزاد شدند تا مطابقت بهتری با دیتاهای خاص پروژه تفکیک زباله داشته باشند اما لایه‌های ابتدایی تا انتها، فریز شده ماندند چون ویژگی‌های عمومی تصاویر را دارند و نیازی به تغییر آن‌ها نیست. این کار به مدل اجازه می‌دهد تا ابتدا روی ویژگی‌های عمومی مانند لبه‌ها و الگوهای پایه تمرکز کند و فقط لایه‌های جدید که اضافه شده اند آموزش ببینند. این تکنیک و استفاده از انتقال یادگیری، توانست کمک کند تا مسیر بهتری برای آموزش مدل، پیگیری شود. در این پژوهش، روش یا الگوریتم مورد استفاده، نسبت به مدل EfficientNet تغییر کرده است. برخلاف مدل ذکر شده، از الگوریتم Adaptive Moment Estimation و یا Adam برای بهینه‌سازی استفاده شد. مدل EfficientNetB0 با RMSProp آموزش دیده است اما در این پروژه از Adam استفاده شده زیرا برای مجموعه داده‌های کوچک‌تر (درقیاس با دیتاهایی که مدل EfficientNetB0 روی آن‌ها آموزش دیده که تصاویر بسیار زیادی هستند) و مسائل سفارشی شده، عملکردی پایدارتر و همگرایی سریع‌تری دارد و برای دیتاهای کوچک‌تر مناسب‌تر است (Kingma & Ba, 2014).

شکل ۳، نشان‌دهنده‌ی درصد تصاویر هر دسته از تصاویر در بین تمام تصاویر مجموعه داده Realwaste پس از حذف دسته متفرقه است.



شکل ۳. درصد وزنی هر دسته از تصاویر، نسبت به کل تصاویر

Figure 3. Weight percentage of each category of images, relative to the total images

همانطور که از شکل ۳ مشخص است، به دلیل عدم توازن در تعداد تصاویر مربوط به دسته‌های مختلف، وزن‌دهی کلاس‌ها به صورت خودکار براساس فراوانی هر کلاس انجام گرفت تا در حین آموزش، مدل نسبت به دسته‌های کم‌نمونه، دقت بیشتری نشان دهد. در این پروژه، از کتابخانه‌هایی در کدنویسی پایتون استفاده گردید؛ دو نمونه از کتابخانه‌های اصلی عبارت اند از: Scikit-learn و TensorFlow، و برنامه‌نویسی در Visual Studio Code انجام گرفت. در مدل تک‌برچسبی، مدل فقط یک برچسب و اطمینان مربوط به همان برچسب را در خروجی نمایش می‌دهد (مدل، برچسبی که بالاترین اطمینان را در آن دارد، خروجی می‌دهد). در این پژوهش داده‌های آموزش و اعتبارسنجی با نسبت ۷۰ درصد به ۳۰ درصد تعیین شدند.

در این پژوهش، هر دو مدل بر پایه‌ی خروجی‌های مستقل sigmoid طراحی شدند؛ به این معنا که برای هر کلاس، یک مقدار بین صفر و یک تولید می‌شود و این مقادیر الزاماً جمعی برابر با یک ندارند (Cheng et al., 2022). این رویکرد در مدل تک‌برچسبی هم برقرار بود و علت آن یکسان‌سازی معماری خروجی دو مدل برای مقایسه‌پذیری ساده‌تر می‌باشد. بنابراین، مدل تک‌برچسبی در زمان پیش‌بینی، تنها برچسبی را که بیشترین مقدار را دارد به عنوان پاسخ نهایی گزارش می‌کند، در حالی که مدل چندبرچسبی همه‌ی برچسب‌هایی را که مقدار خروجی آن‌ها از آستانه‌ی تعیین شده (۰/۲۵) بیشتر باشد نمایش می‌دهد.

یافته‌ها

در این پژوهش، با بهره‌گیری از توانمندی‌های یادگیری عمیق و بینایی ماشین، گامی موثر در جهت توسعه‌ی سامانه‌ای هوشمند برای طبقه‌بندی و تفکیک خودکار زباله‌ها برداشته شد. استفاده از معماری کارآمد EfficientNetB0 و بهره‌گیری از تکنیک‌هایی مانند افزایش مصنوعی داده‌ها، فریز کردن لایه‌ها و بهینه‌سازی الگوریتم آموزش باعث شد که مدل نهایی، عملکردی مناسب و قابل قبول در دسته‌بندی تصاویر زباله‌ها از خود نشان دهد که نتایج حاصل در قالب شکل، آورده شده است.

ساخت دو مدل مجزا برای طبقه‌بندی تک‌برچسبی و چندبرچسبی، امکان مقایسه و بررسی کارایی رویکردهای مختلف در تفکیک زباله را فراهم ساخت. در مدل تک‌برچسبی تفاوت بین خطای آموزشی^۱ و خطای اعتبارسنجی^۲ در طول آموزش مدل، مقداری افزایش یافت؛ این نشان‌دهنده‌ی این است که احتمال بیش‌برازش خفیف وجود دارد اما به اندازه‌ای نیست که به مدل آسیب جدی بزند. در مدل چندبرچسبی، فاصله‌ی خطای آموزشی و اعتبارسنجی حین آموزش مدل، متعادل‌تر و مناسب بود که مدل احتمالاً نه کم‌برازش^۳ است و نه بیش‌برازش شده که اتفاق خوبی است. افزایش مصنوعی بیشتر داده‌ها و استفاده از تصاویر ترکیبی، در این موضوع تاثیرگذار بوده اند. همچنین کاهش تعداد دوره‌های آموزش و یا اپک‌ها در جلوگیری از بیش‌برازش شدن مدل چندبرچسبی تاثیرگذار بوده است؛ چراکه پس از تشخیص بیش‌برازش خفیف در مدل تک‌برچسبی، در آموزش مدل چندبرچسبی، تعداد اپک‌ها کاهش داده شد. با اینکه مدل چندبرچسبی از نظر بیش‌برازش شدن محافظت بیشتری داشت اما همانطور که گفته شده، سطح بیش‌برازش شدن مدل تک‌برچسبی نیز قابل قبول است.

مدل‌های تک‌برچسبی و چندبرچسبی با تصاویری جدید که مدل‌ها در زمان آموزش، آن‌ها را ندیده اند، تست شدند (این تصاویر از اینترنت دانلود شدند). نتایج حاصل از ارزیابی نهایی مدل‌ها روی تصاویر جدید، نشان‌دهنده نقاط قوت

¹ training loss

² validation loss

³ Underfit

و ضعف مدل‌ها است. با استفاده از این نتایج می‌توان مشخص کرد که مدل‌های فعلی در چه مواردی کاربرد دارند و جهت چه استفاده‌هایی نیاز به ارتقاء مدل‌های فعلی وجود دارد.



شکل ۴. تصویر دسته پسماندهای کاغذی

Figure 4. Image of paper waste pile

اولین تصویر دانلود شده از اینترنت که مدل‌ها روی آن تست شدند یعنی تصویر شکل ۴، متعلق به دسته‌ی پسماندهای کاغذی است که خروجی چاپ شده مدل چندبرچسبی در ترمینال^۱ نرم‌افزار به این صورت بود:

Predicted Labels²: Paper: 0.76

مدل چندبرچسبی با اطمینان ۰/۷۶ از ۱ یا ۷۶ درصدی، پسماند کاغذی تشخیص داد که صحیح است. مدل برای این تصویر تنها یک برچسب تولید کرد که نشان می‌دهد اطمینان مدل برای تعلق گرفتن این تصویر به سایر دسته‌ها، کمتر از حد آستانه است که در این پروژه، ۲۵ درصد تعیین شده است. با توجه به اینکه تصویر دقیقاً در دسته‌ی پسماندهای کاغذی قرار دارد، مدل به درستی عمل کرده است. خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Paper (0.84)

مدل تک‌برچسبی با اطمینان ۰/۸۴ از ۱ یا ۸۴ درصدی، پسماند کاغذی تشخیص داد که صحیح است. در مورد این تصویر مشاهده می‌شود که مدل تک‌برچسبی بهتر عمل کرده است (۸ درصد اطمینان بالاتر دارد).



شکل ۵. تصویر دسته پسماندهای نساجی/پارچه‌ای

Figure 5. Image of textile/fabric waste category

دومین تصویر دانلود شده از اینترنت که مدل روی آن تست شد یعنی شکل ۵، متعلق به دسته‌ی پسماندهای نساجی/پارچه‌ای است که خروجی چاپ شده مدل چندبرچسبی در ترمینال به این صورت بود:

Predicted Labels: Textile Trash: 0.87

¹ Terminal

² برچسب‌های پیش‌بینی شده

مدل چندبرچسبی با اطمینان ۰/۸۷ از ۱ یا ۸۷ درصدی، پسماند نساجی/پارچه‌ای تشخیص داد که صحیح است. مدل برای این تصویر فقط یک برچسب تولید کرد. خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Textile Trash (0.80)

مدل تک‌برچسبی با اطمینان ۰/۸۰ از ۱ یا ۸۰ درصدی، پسماند نساجی/پارچه‌ای تشخیص داد که صحیح است. درمورد این تصویر مدل چندبرچسبی با ۸۷ درصد اطمینان، بهتر از مدل تک‌برچسبی با اطمینان ۸۰ درصدی عمل کرد، اما هر دو مدل برچسب صحیح تولید کردند.



شکل ۶. تصویر دسته پسماندهای آلی غذایی
Figure 6. Image of organic food waste category

سومین تصویر دالود شده از اینترنت که مدل روی آن تست شد یعنی شکل ۶، متعلق به دسته‌ی پسماندهای آلی غذایی است که خروجی چاپ شده مدل چندبرچسبی در ترمینال به این صورت بود:

Predicted Labels: Food Organics: 0.93

Predicted Labels: Textile Trash: 0.31

مدل دو برچسب تولید کرد؛ با اطمینان ۰/۹۳ از ۱ یا ۹۳ درصدی، پسماند آلی غذایی و با اطمینان ۰/۳۱ از ۱ یا ۳۱ درصدی، پسماند نساجی/پارچه‌ای پیش‌بینی شدند. از آنجا که مدل با اطمینان سه برابر بالاتر، تصویر را پسماند آلی غذایی تشخیص داده پس می‌توان برچسب نسبت داده شده از مدل برای تصویر را همان Food Organics در نظر گرفت که صحیح است. همچنین در صورت تمایل می‌توان کدنویسی کرد و این شرط را تعیین کرد که به عنوان مثال اگر اطمینان مدل بالای ۶۵ درصد بود، برچسب نمایش داده شود. با انجام این کار، زمانی که اطمینان مدل زیاد نیست، برچسب مدل نمایش داده نمی‌شود و این عدد (۶۵ درصد) را نیز می‌توان به دلخواه در کد تست مدل، تغییر داد. خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Food Organics (0.87)

مدل تک‌برچسبی با اطمینان ۰/۸۷ از ۱ یا ۸۷ درصدی، پسماند آلی غذایی تشخیص داد که صحیح است. مدل چندبرچسبی با اطمینان ۹۳ درصدی، درمورد تشخیص برچسب صحیح این تصویر، بهتر عمل کرده است.



شکل ۷. تصویر دسته پسماندهای پلاستیکی

Figure 7. Image of plastic waste pile

چهارمین تصویر دانلود شده از اینترنت که مدل روی آن تست شد یعنی شکل ۷، متعلق به دسته‌ی پسماندهای پلاستیکی است که خروجی چاپ شده مدل چندبرچسبی در ترمینال به این صورت بود:

Predicted Labels: Plastic: 0.94

مدل چندبرچسبی با اطمینان ۰/۹۴ از ۱ یا ۹۴ درصدی، پسماند پلاستیکی تشخیص داد که صحیح است. مدل برای این تصویر فقط یک برچسب تولید کرد. خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Plastic (0.91)

مدل تک‌برچسبی با اطمینان ۰/۹۱ از ۱ یا ۹۱ درصدی، پسماند پلاستیکی تشخیص داد که صحیح است. مدل چندبرچسبی با اطمینان ۹۴ درصدی، درمورد تشخیص برچسب صحیح این تصویر، بهتر عمل کرده است. با مقایسه‌ی عملکرد هر دو مدل در پیش‌بینی برچسب تصاویر ۴ تا ۷، می‌توان نتیجه گرفت که مدل چندبرچسبی در ۳ مورد از ۴ مورد، با اطمینان بیشتری دسته‌ی صحیح را تشخیص داده است و همچنین متوسط اطمینان مدل چندبرچسبی در تشخیص دسته‌ی صحیح ۰/۸۷۵ از ۱ بوده است و این عدد برای مدل تک‌برچسبی، ۰/۸۵۵ از ۱ بوده است.



شکل ۸. تصویر دسته متفرقه

Figure 8. Miscellaneous category image

پنجمین تصویر دانلود شده از اینترنت که مدل روی آن تست شد یعنی شکل ۸، متعلق به هیچ‌کدام از ۸ دسته اصلی که مدل روی تصاویر آن‌ها آموزش دیده است نیست که خروجی چاپ شده مدل چندبرچسبی در ترمینال به این صورت بود که یعنی: هیچ برچسبی پیش‌بینی نشد (همه‌ی اطمینان‌ها کمتر از آستانه‌ی تعیین شده اند):

No labels predicted (all confidence below threshold).

مدل برچسبی اختصاص نداد؛ چراکه ویژگی‌های موجود در تصویر، با ویژگی‌های تصاویری که مدل روی آن‌ها آموزش دیده (۸ دسته)، بسیار متفاوت است و مدل به اطمینان قابل قبولی جهت ارائه برچسب نرسید. حداقل آستانه تعیین شده برای مدل جهت چاپ برچسب، اطمینان ۰/۲۵ از ۱ یا ۲۵ درصدی است.

خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Metal (0.29)

مدل تک‌برچسبی اطمینان پایینی روی این تصویر دارد (۰/۲۹ از ۱ یا ۲۹ درصدی). اما از آنجا که حداقل اطمینان لازم برای چاپ برچسب ۲۵ درصد تعیین شده است (که می‌توان آن را تغییر داد)، مدل تک‌برچسبی، برچسب پسماند فلزی را تولید کرده است، اما مدل چندبرچسبی از آنجا که هیچ برچسبی تولید نکرده است، بنابراین اطمینانی پایین‌تر از ۲۵ درصد برای این تصویر در تمام دسته‌ها داشته است و از آنجا که این تصویر در هیچ دسته‌ای قرار نمی‌گیرد، بنابراین می‌توان گفت مدل چندبرچسبی بهتر عمل کرده است.



شکل ۹. تصویر دسته تصاویر پیچیده (حضور هم‌زمان چند نوع پسماند)

Figure 9. Image of a complex image cluster (simultaneous presence of multiple types of waste)

ششمین تصویر دانلود شده از اینترنت که مدل روی آن تست شد یعنی شکل ۹، تصویری است که از ترکیب پنج نوع مختلف از تصاویر پسماند که مدل‌ها روی آن‌ها آموزش دیده اند ایجاد شده است، که تصویری پیچیده به حساب می‌آید. هدف از ارائه این تصویر به مدل‌ها این است که توانایی آن‌ها روی تصاویر پیچیده سنجیده شود. خروجی چاپ شده مدل چندبرچسبی در ترمینال به این صورت بود:

Predicted Labels: Food Organics: 0.35

Predicted Labels: Glass: 0.30

مشاهده می‌شود که مدل نتوانست تمام برچسب‌ها را به درستی پیش‌بینی کند و همچنین مدل اطمینان‌های پایینی دارد (۰/۳۵ از ۱ یا ۳۵ درصد و ۰/۳۰ از ۱ یا ۳۰ درصد)، البته برچسب‌هایی که مدل پیش‌بینی کرده است، صحیح هستند (با اطمینان ۳۰ درصدی پسماند شیشه‌ای و با اطمینان ۳۵ درصدی پسماند آلی غذایی) اما این یک خروجی ناقص است. این تصویر دقیقاً نقطه ضعف مدل‌های فعلی را بیان می‌کند. در صورتی که هدف، تحلیل تصاویر شلوغ که دربرگیرنده چندین دسته مختلف هستند باشد، استفاده از مدل‌های تشخیص اشیاء مثل YOLO یکی از راه‌حل‌های مقابله با این مشکل است.

خروجی مدل تک‌برچسبی در ترمینال نرم‌افزار:

Predicted Class: Food Organics (0.28)

مدل تک‌برچسبی با اطمینان ۰/۲۸ از ۱ یا ۲۸ درصدی، پسماند آلی غذایی تشخیص داد. از آنجایی که مدل تک‌برچسبی تنها یک برچسب تولید می‌کند، مدل تصویر را پسماند آلی غذایی تشخیص داد که یکی از پنج تصویر

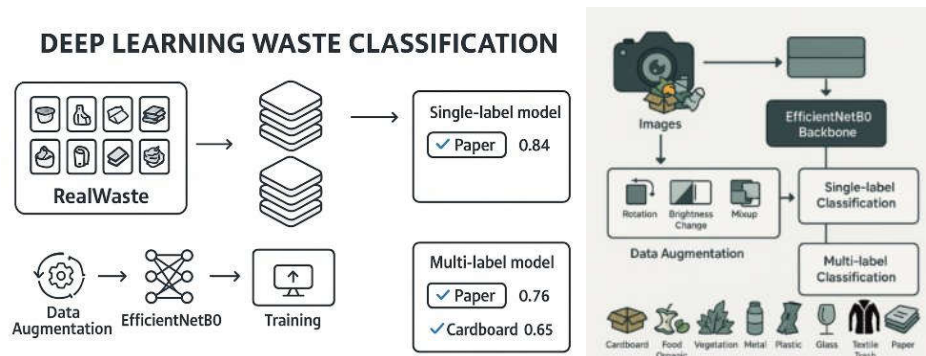
شکل ۵، پسماند آلی غذایی است اما به جهت پیچیدگی تصویر، مدل دقت پایینی دارد. همچنین مشاهده می‌شود که مدل چندبرچسبی نیز برچسب را تولید کرده است اما با دقت بالاتری نسبت به مدل تک‌برچسبی (دقت ۳۵ درصدی) بنابراین درمورد این تصویر مدل چندبرچسبی بهتر عمل کرده است.

مقایسه تجربی نمونه‌های آزمون که شرح داده شد، نشان داد که مدل چندبرچسبی در بسیاری از موارد با اطمینان مناسب‌تری برچسب‌های درست را گزارش کرده است. یکی از دلایل محتمل این رفتار، کاهش تعداد دورها در آموزش مدل چندبرچسبی است؛ زیرا مدل تک‌برچسبی، نشانه‌هایی از بیش‌برازش خفیف را نشان می‌داد، در حالی که در مدل چندبرچسبی با کاهش تعداد دورهای آموزش، این مشکل تا حد زیادی کنترل شد. در نتیجه به عنوان یکی از دلایل، بهبود مشاهده شده در مدل چندبرچسبی را می‌توان تا حدی به کاهش بیش‌برازش نسبت داد. همچنین برتری مدل چندبرچسبی در این مقاله باید به صورت یک مشاهده‌ی تجربی تفسیر شود. مشاهده‌ی تجربی نشان از آن داشت که مدل چندبرچسبی در تحلیل تصاویر با یک شیء، از مدل تک‌برچسبی بهتر عمل کرد اما به دلیل محدودیت‌های معماری EfficientNet در تحلیل تصاویر با چند شیء، مدل چندبرچسبی نتوانست در تحلیل تصویر با چندین پسماند، اطمینان قابل قبولی ارائه دهد، هرچند در این تحقیق، مدل چندبرچسبی در تحلیل تصاویر پیچیده، بهتر از مدل تک‌برچسبی عمل کرد. به همین منظور از نتایج پژوهش می‌توان برداشت کرد که جهت تحلیل تصاویر چندشیئی، بهتر است از انواع دیگری از مدل‌ها کمک گرفت که YOLO پیشنهاد می‌شود.

بحث

در این پژوهش، عملکرد مدل‌ها به صورت تجربی و با اعمال آن‌ها بر چند تصویر جدید دانلود شده از اینترنت در حالت‌های مختلف، مقایسه و ارزیابی شد. با این حال، جداسازی درصد قابل قبول و مناسبی از داده‌ها پیش از آموزش مدل و سپس ارزیابی مدل نهایی بر روی این داده‌ها که در فرایند آموزش توسط مدل مشاهده نشده اند، همراه با گزارش معیارهایی مانند دقت، precision، یادآوری و F1-score، می‌تواند به شناخت دقیق‌تر عملکرد مدل‌ها کمک کند. در پژوهش حاضر، برای امکان‌پذیر شدن مقایسه‌ی تجربی مدل‌ها در تشخیص دسته‌ی صحیح، خروجی مدل تک‌برچسبی نیز همانند مدل چندبرچسبی به صورت sigmoid تعریف شد و در مرحله‌ی پیش‌بینی، مدل تک‌برچسبی برچسب دارای بیشترین مقدار را به عنوان خروجی نهایی در نظر گرفت. با این حال، ارزیابی مدل‌ها با مجموعه داده‌ی آزمون مستقل، تصاویر بیشتر و معیارهای کمی متنوع‌تر، می‌تواند در پژوهش‌های آتی دنبال شود.

استفاده از مدل‌های تشخیص اشیاء مثل YOLO می‌تواند مدل را برای تحلیل تصاویر پیچیده‌تر و استفاده در محل‌های دفن زباله در دنیای واقعی، قابل اعتمادتر و مناسب‌تر کند. در این پژوهش از مدل EfficientNetB0 به عنوان مدل پایه استفاده شد و مدل فعلی، علی‌رغم تمام نقطه قوت‌هایش، در موقعیت‌هایی که تصویر دارای بیش از یک نوع زباله است یا تصویر پیچیده است، ممکن است کارآیی لازم را در مقایسه با مدل‌هایی مثل YOLO نداشته باشد. استفاده از مدل‌های YOLO می‌تواند این مشکل را به خوبی مدیریت کند. مدل‌های تشخیص شیء، هم موقعیت شیء و هم نوع آن را در تصاویر پیچیده با چندین نوع پسماند در یک تصویر مشخص می‌کنند. در این پژوهش محدودیت مدل‌های EfficientNet در تحلیل تصاویر چندشیئی دیده شد اما استفاده از YOLO انجام نشد. به کار بردن YOLO در پژوهش‌های آتی، توصیه می‌گردد.



شکل ۱۰. مدل گرافیکی پروژه

Figure 10. Graphical model of the project

شکل ۱۰، مدل گرافیکی پروژه شامل افزایش مصنوعی داده‌ها، معماری مدل، طبقه‌های پیش‌بینی شده توسط مدل و انواع مدل‌ها را نشان می‌دهد.

نتیجه‌گیری

استفاده از یادگیری ماشین و به خصوص یادگیری عمیق باعث می‌شود بتوان مسائل پیچیده را با هزینه‌ی کمتر و سرعت و دقت بالاتر انجام داد و محیط زیست یکی از حوزه‌هایی است که می‌تواند از این پتانسیل به خوبی بهره‌برد. در این پژوهش، با بهره‌گیری از توانمندی‌های یادگیری عمیق و بینایی ماشین، گامی موثر در جهت توسعه سامانه‌ای هوشمند برای طبقه‌بندی و تفکیک خودکار زباله‌ها در هشت دسته برداشته شد. استفاده از معماری کارآمد EfficientNetB0 و بهره‌گیری از تکنیک‌هایی مانند افزایش مصنوعی داده‌ها، فریز کردن لایه‌ها، وزن‌دهی کلاس‌ها و بهینه‌سازی الگوریتم آموزش با Adam منجر به این شد که مدل‌های نهایی، عملکردی مناسب و قابل قبول در دسته‌بندی تصاویر زباله‌ها از خود نشان دهند. ساخت دو مدل مجزا برای طبقه‌بندی تک‌برچسبی و چندبرچسبی، امکان مقایسه و بررسی کارایی رویکردهای مختلف در تفکیک زباله را فراهم ساخت.

نتایج حاصل از ارزیابی مدل‌ها روی تصاویر جدید نشان داد که در این تحقیق هر دو مدل در طبقه‌بندی تصاویر، عملکرد مناسبی داشتند، اما مدل چندبرچسبی در مجموع نسبت به مدل تک‌برچسبی، دقیق‌تر عمل کرد. مدل چندبرچسبی همچنین در تشخیص تصویر نامرتب با دسته‌های آموزش دیده و تصویر پیچیده که شامل چندین نوع پسماند بود، عملکرد بهتری از خود نشان داد. مدل تک‌برچسبی در تحلیل تصاویر با یک نوع پسماند گاهی اطمینان بالاتری ارائه داد اما مدل چندبرچسبی به دلیل مدیریت بهتر آموزش مدل، در مجموع در آزمون تجربی موفق‌تر عمل کرد. به طور کلی رویکرد چندبرچسبی به دلیل انعطاف‌پذیری بیشتر در تشخیص هم‌زمان چندین کلاس، برای کاربردهای واقعی در خطوط بازیافت که با تنوع بالای پسماندها مواجه هستند، گزینه‌ی مناسب‌تری محسوب می‌شود که البته معماری EfficientNet، انتخاب مطمئن و یا بهترین انتخاب برای تحلیل تصاویر با چندین نوع پسماند در تصاویر پیچیده نمی‌باشد.

از منظر نظری، این پژوهش با مقایسه‌ی عملکرد مدل‌های تک‌برچسبی و چندبرچسبی در حوزه تفکیک پسماند، به غنای ادبیات موجود در زمینه‌ی کاربرد یادگیری عمیق در مدیریت محیط زیست کمک کرده است. از منظر عملی، مدل‌های توسعه یافته می‌توانند ضمن ارزیابی‌های لازم و بررسی انطباق از نظر ماهیت و نوع میان تصاویری که مدل روی آن‌ها آموزش دیده با تصاویری که مدل قرار است آن‌ها را تحلیل کند؛ در خطوط بازیافت و سامانه‌های هوشمند مدیریت پسماند مورد استفاده قرار گیرند و به کاهش هزینه‌ها، تسریع در انجام امور و کاهش خطای انسانی کمک

کنند. استقرار مدل‌های پردازش تصویر در سامانه‌های هوشمند مدیریت پسماند می‌تواند گامی موثر در جهت تحقق اهداف توسعه پایدار و مدیریت بهینه‌ی منابع محسوب شود.

References

- Aghdam, F. Z., Hasanlou, M., & Dehghanijabbarlou, M. (2024). Quantifying urban air quality through multispectral satellite imagery and Google earth Engine. *Journal of Atmospheric and Solar-Terrestrial Physics*, 261, 106301. DOI: 10.1016/j.jastp.2024.106301
- Ahmad, G., Aleem, F. M., Alyas, T., Abbas, Q., Nawaz, W., Ghazal, T. M., ... & Ibrahim, A. M. (2025). Intelligent waste sorting for urban sustainability using deep learning. *Scientific reports*, 15(1), 27078. DOI: 10.1038/s41598-025-08461-w
- Ali, H., Shifa, N., Benlamri, R., Farooque, A. A., & Yaqub, R. (2025). A fine tuned EfficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases. *Scientific Reports*, 15(1), 25732. DOI: 10.1038/s41598-025-04479-2
- Alotaibi, E., & Nassif, N. (2024). Artificial intelligence in environmental monitoring: in-depth analysis. *Discover Artificial Intelligence*, 4(1), 84. DOI: 10.1007/s44163-024-00198-1
- Arishi, A. (2025). Real-time household waste detection and classification for sustainable recycling: A deep learning approach. *Sustainability*, 17(5), 1902. DOI: 10.3390/su17051902
- Bai, T., Wang, L., Yin, D., Sun, K., Chen, Y., Li, W., & Li, D. (2023). Deep learning for change detection in remote sensing: a review. *Geo-spatial Information Science*, 26(3), 262-288. <https://doi.org/10.1080/10095020.2022.2085633>
- Bayatzadeh, M., Hosseini Rahman, S. A., Karbalai Hasani, H., & Sheikhi, M. (2025). *Dry waste detection based on an intelligent convolutional neural network system in the Arak Municipal Waste Management Organization Paya Shahr*, 7(77). [In Persian]. <https://civilica.com/doc/2333531>
- Biglarijoo, N., Shams, A., & Hadipour, V. (2025). The use of machine learning techniques in water and wastewater treatment processes: opportunities and challenges. *Numerical Methods in Civil Engineering*, 10(2), 86-101. DOI: 10.66224/NMCE.2511.1106
- Chen, L., Li, S., Bai, Q., Yang, J., Jiang, S., & Miao, Y. (2021). Review of image classification algorithms based on convolutional neural networks. *Remote Sensing*, 13(22), 4712. DOI: 10.3390/rs13224712
- Chen, Y. C. (2018). Effects of urbanization on municipal solid waste composition. *Waste management*, 79, 828-836. DOI: 10.1016/j.wasman.2018.04.017
- Cheng, C. S., Chen, P. W., & Ho, Y. (2022). Control chart concurrent pattern classification using multi-label convolutional neural networks. *Applied Sciences*, 12(2), 787. <https://www.mdpi.com/2076-3417/12/2/787>
- Dawar, I., Srivastava, A., Singal, M., Dhyani, N., & Rastogi, S. (2025). A systematic literature review on municipal solid waste management using machine learning and deep learning. *Artificial Intelligence Review*, 58(6), 183. DOI: 10.1007/s10462-025-11196-9
- De Mauro, A., Sestino, A., & Bacconi, A. (2022). Machine learning and artificial intelligence use in marketing: a general taxonomy. *Italian Journal of Marketing*, 2022(4), 439-457. DOI: 10.1007/s43039-022-00057-w
- Dhiman, R., Miteff, S., Wang, Y., Ma, S. C., Amirikas, R., & Fabian, B. (2024). Artificial intelligence and sustainability—A review. *Analytics*, 3(1), 140-164. DOI: 10.3390/analytics3010008
- Fang, B., Yu, J., Chen, Z., Osman, A. I., Farghali, M., Ihara, I., ... & Yap, P. S. (2023). Artificial intelligence for waste management in smart cities: a review. *Environmental Chemistry Letters*, 21(4), 1959-1989. <https://doi.org/10.1007/S10311-023-01604-3>
- Himeur, Y., Rimal, B., Tiwary, A., & Amira, A. (2022). Using artificial intelligence and data fusion for environmental monitoring: A review and future perspectives. *Information Fusion*, 86, 44-75. DOI: 10.1016/j.inffus.2022.06.003
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260. DOI: 10.1126/science.aaa8415

- Kaza, S., Yao, L., Bhada-Tata, P., & Van Woerden, F. (2018). *What a waste 2.0: a global snapshot of solid waste management to 2050*. World Bank Publications.
- Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access*, 12, 94250-94295. DOI: 10.1109/ACCESS.2024.3425166
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. <https://doi.org/10.48550/arXiv.1412.6980>
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS journal of photogrammetry and remote sensing*, 152, 166-177. DOI: 10.1016/j.isprsjprs.2019.04.015
- Mehadjbia, A., & Slaoui-Hasnaoui, F. (2024). Real-time waste detection based on YOLOv8. In *2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC)* (pp. 1-6). IEEE. DOI: 10.1109/INTCEC61833.2024.10603365
- Mohammadzadeh, M., Delavar, M. R., & Rabiei Dastjerdi, H. (2026). Mapping Urban Garden Destruction in Tehran Using Landsat Imagery and Machine Learning. *ISPRS Annals of the Photogrammetry. Remote Sensing and Spatial Information Sciences*, 10, 533-540. <https://doi.org/10.5194/isprs-annals-X-4-W8-2025-533-2026>
- Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, 2(6), 1-20. DOI: 10.1007/s42979-021-00815-1
- Sayem, F. R., Islam, M. S. B., Naznine, M., Nashbat, M., Hasan-Zia, M., Kunju, A. K. A., ... & Chowdhury, M. E. (2025). Enhancing waste sorting and recycling efficiency: robust deep learning-based approach for classification and detection. *Neural Computing and Applications*, 37(6), 4567-4583. DOI: 10.1007/s00521-024-10855-2
- Single, S., Iranmanesh, S., & Raad, R. (2023). Realwaste: a novel real-life data set for landfill waste classification using deep learning. *Information*, 14(12), 633. DOI: 10.3390/info14120633
- Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR. DOI: 10.48550/arXiv.1905.11946
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, 11(1), 233. DOI: 10.48550/arXiv.1905.00501
- Wu, J., & He, J. (2025). Trustworthy Transfer Learning: A Survey. *Journal of Artificial Intelligence Research*, 84. <https://doi.org/10.1613/jair.1.176021>
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., ... & Zhang, L. (2020). Deep learning in environmental remote sensing: Achievements and challenges. *Remote sensing of Environment*, 241, 111716. DOI: 10.1016/j.rse.2020.111716
- Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*, 5(4), 8-36. DOI: 10.1109/MGRS.2017.2762307
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76. DOI: 10.1109/JPROC.2020.3004555