

Clasificación automática de residuos aprovechables y no aprovechables en la Pontificia Universidad Javeriana Cali usando Deep Learning.

Juan Esteban Torres Tamayo¹, Janner Arley Rosero Mora¹

¹Pontificia Universidad Javeriana, Cali, Valle del Cauca, Colombia

Este documento propone la implementación de un algoritmo de clasificación automática de residuos aprovechables y no aprovechables dentro de las instalaciones del campus de la Pontificia Universidad Javeriana Cali haciendo uso de técnicas de Deep Learning. En este se presenta un benchmarking entre 6 modelos implementados haciendo uso de métodos de Transfer Learning, Fine Tuning y Feature Extractor, donde cada modelo es evaluado con 6 métricas de clasificación. Para el entrenamiento y validación de estos se hizo uso de un dataset propio compuesto de 2015 fotografías de residuos, 1000 para aprovechables y 1015 para no aprovechables, este dataset fue elaborado dentro de las instalaciones de la Unidad Central de Almacenamiento de la Pontificia Universidad Javeriana Cali con el fin de obtener un mapeo óptimo de los residuos generados dentro de la institución por parte de los colaboradores y estudiantes. Después de los procesos de preprocesamiento, optimización, entrenamiento y validación se logró determinar que de los modelos probados el que mejor resultados brindó fue la arquitectura MobileNetV3 con Fine Tuning que alcanzó un Accuracy de 0.938 con el dataset original y un 0.9433 al aplicar aumento de datos con transformación radiométrica.

Introducción

La correcta clasificación y el procesamiento de los residuos es una problemática de magnitud mundial que viene perjudicando el medio ambiente desde hace décadas, tanto así que la organización de las Naciones Unidas lo tiene contemplado dentro de los Objetivos de Desarrollo Sostenible (ODS). Haciendo énfasis en Colombia, para minimizar las consecuencias y normatizar estos procedimientos se expidió la resolución 2184 de 2019 la cual dicta que existen 3 clasificaciones para los residuos: Aprovechable (Blanco), No Aprovechable (Negro) y Orgánico (Verde). La Pontificia Universidad Javeriana Cali cuenta con el Plan de Gestión Integral de Residuos (PGIR), en el cual dicta el tratamiento que le brindan a los residuos generados dentro de sus instalaciones, al analizar a profundidad este proceso se encuentra la mayoría de los residuos pasan por una clasificación manual dentro de la Unidad Central de Almacenamiento (UCA) donde un operario abre las bolsas recolectadas y determina la clasificación de los residuos, los cuales terminan en un relleno sanitario (No Aprovechables) o son procesados y reciclados por empresas especializadas (Aprovechables).

Con el fin de brindar el apoyo de una herramienta tecnológica que cuente con la capacidad de clasificar automáticamente los residuos aprovechables y no aprovechables que llegan a la UCA, se implementan modelos que parten de técnicas de Deep Learning y visión artificial con el fin de

determinar el modelo que mejores resultados brinda después de un benchmarking que evalúa cada uno con seis métricas de clasificación. Para dicho proceso es necesario el uso de un dataset óptimo, y eso por ello que se construye un ambiente de luminosidad controlada que permita la captura fotográfica de distintos residuos que posteriormente son clasificados entre dos etiquetas: Aprovechable y No Aprovechable.

Fundamentación teórica

Para el desarrollo de este proyecto se hizo uso de varios campos tecnológicos que desde sus respectivas áreas aportan para el cumplimiento del objetivo. Si bien todos cumplen un papel fundamental, el principal fundamento teórico es la clasificación de los residuos dentro del margen establecido por la resolución 2184 de 2019 [1], este dicta que existen tres tipos de residuos:

- Residuos aprovechables: Son residuos que por sus características materiales pueden ser procesados ser reintroducidos en la cadena productora [2]. Entre ellos se pueden encontrar los plásticos, cartón, vidrios y/o metales [1].
- Residuos no aprovechables: Son residuos que dadas sus características físicas no cuentan con la capacidad para ser aprovechados de nuevo [2]. Entre ellos se pueden encontrar papeles metalizados, servilletas, papel o cartón contaminado con restos de comida, entre otros.

En base a esta clasificación se procede ahora a indagar el como lograr que un algoritmo logre aprender a clasificar residuos a partir de una fotografía, y es así como se apoya del campo del Deep Learning, el cual es un subcampo del aprendizaje de máquina que busca emular el humano con el fin de dar poder decisional a una máquina. Una de las características que lo hace tan versátil es que a diferencia del aprendizaje de máquina no se requiere de reglas previamente programadas pues el propio sistema en sí es capaz de “Aprender” tras una sesión de entrenamiento. Otra característica relevante dentro de este método es que se compone de redes neuronales entrelazadas que procesan la información para retornar una salida [3].

Si bien el Deep Learning aporta las técnicas suficientes para el desarrollo del proyecto se decidió ir más allá y hacer uso del Transfer Learning, el cual es un método de aprendizaje de máquina en el que un modelo desarrollado para una tarea se reutiliza como punto de partida para un modelo que cuenta con una tarea distinta. Este método ya es un enfoque popular en donde se utilizan modelos preentrenados en tareas de visión artificial o procesamiento del lenguaje natural y gracias a este entrenamiento previo es que el segundo modelo puede conseguir aprovechar este procesamiento para obtener buenos resultados [4].

Haciendo hincapié en el proceso de benchmarking, para llevar a cabo este es necesario un conjunto de métricas que permitan evaluar el rendimiento de cada modelo, dadas las características del proyecto se seleccionan seis métricas de clasificación:

- Accuracy: Esta métrica consiste en la relación entre el número de muestras clasificadas correctamente y el número total de muestras. Sin embargo, esta métrica no puede ser considerada adecuada cuando el dataset en uso está desbalanceado [5].
- MCC (Matthews Correlation Coefficient): Una métrica que ha tomado gran acogida y permite una solución a los problemas de clases desbalanceadas es el Coeficiente de Correlación de Matthews [5].

- **F1 score:** Esta métrica es usualmente usada tanto en problemas binarios como multiclases y es altamente usada dado a su invariabilidad [5].
- **Matriz de confusión:** Las matrices de confusión son la forma más usual y estándar de informar sobre la exactitud temática de productos derivados de la clasificación de datos procedentes de imágenes [6].
- **Curva ROC:** La curva ROC es una gráfica que presenta el rendimiento de un algoritmo de clasificación en todos los umbrales [7].
- **Área bajo la curva ROC (AUC):** El AUC mide el área bidimensional bajo la curva ROC tomando valores entre 0 y 1, donde 0 se interpreta como un clasificador con predicciones 100% incorrectas y 1 como un clasificador con predicciones 100% correctas [8].

Una vez se cuenta con los fundamentos teóricos presentados anteriormente se parte ahora a la construcción de un dataset óptimo para el proceso de entrenamiento y validación, para ello se hace la captura de 2015 fotografías, 1000 aprovechables y 1015 no aprovechables, todas dentro de un ambiente de luminosidad controlada construido con paredes de acrílico recubiertas de papel Contac blanco y una cinta led de color blanco. Sus dimensiones son presentadas a continuación.

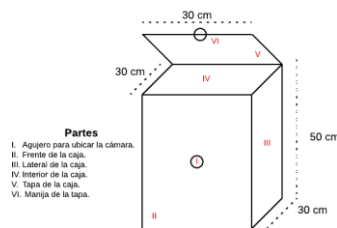


Figura 1. Dimensiones y diseño del ambiente controlado

Resultados

Después de contar con una sólida base fundamental se implementan seis modelos de Deep Learning distintos los cuales son presentados en la tabla 1.

Tabla 1. Modelos usados en el benchmarking

Modelo	Método usado
InceptionV1	Arquitectura InceptionV1 con Fine Tuning.
ResNet50V2	Arquitectura ResNetV2 con Fine Tuning.
MobileNetV3	Arquitectura MobileNetV3 con Fine Tuning.
VGG19	Arquitectura VGG19 con Fine Tuning.
MobileNetV3 + SVM	Arquitectura MobileNetV3 con Feature Extractor y a una máquina de soporte vectorial (SVM).
VGG19 + SVM	Arquitectura VGG19 con Feature Extractor y a una máquina de soporte vectorial (SVM).

Para su entrenamiento y validación se adoptó el método de validación Hold Out con una partición de 80% para entrenamiento y 20% para validación. De allí se procede con el siguiente plan de pruebas:

- **Preprocesamiento de imágenes:** Se cargan todas las fotografías del dataset y son reescaladas a una dimensión de 250x250 píxeles.

- Partición del dataset: El dataset es particionado según el método Hold Out en 1612 fotos para entrenamiento y 403 para validación.
- Entrenamiento: Cada modelo es entrenado 10 veces, donde en cada entrenamiento la partición del dataset varía de acuerdo con una semilla pseudoaleatoria que es almacenada para replicar los resultados.
- Análisis de rendimiento: Finalizado cada entrenamiento por modelo, este es evaluado con 6 métricas: Accuracy, Recall, Specificity, AUC, F1 Score y MCC.

Al finalizar este plan de pruebas se cuenta con el registro de 10 entrenamientos por modelo, de los cuales se almacenan la semilla de partición y cada valor de métrica por entrenamiento. Con esta información se procede entonces a presentar la media y la desviación estándar de cada resultado por modelo en la tabla 2.

Tabla 2. Resultados de entrenamiento

Modelo	Accuracy		Recall		Specificity		AUC		F1 Score		MCC	
	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar	Media	Desviación estándar
InceptionV1	0,8536	0,0314	0,8357	0,1014	0,8710	0,0508	0,8534	0,0329	0,8462	0,0464	0,7135	0,0563
ResNet50V2	0,8625	0,0063	0,8422	0,0196	0,8838	0,0246	0,9212	0,0037	0,8623	0,0061	0,7266	0,0130
MobileNetV3	0,9199	0,0111	0,9199	0,0155	0,9123	0,0144	0,9670	0,0085	0,9213	0,0101	0,8396	0,0224
MobileNetV3+SVM	0,8872	0,0182	0,8872	0,0314	0,8957	0,0173	0,9499	0,0119	0,8883	0,0181	0,7752	0,0360
VGG19	0,8997	0,0200	0,8997	0,0338	0,8934	0,0243	0,9359	0,0178	0,8997	0,0189	0,8105	0,0496
VGG19+SVM	0,8936	0,0180	0,8936	0,0264	0,9040	0,0219	0,9580	0,0093	0,8938	0,0172	0,7875	0,0362

Discusión

Al analizar la información de la tabla 2 se puede determinar que el modelo que presenta mejores resultados en todas las métricas es el modelo MobileNetV3 con método de Fine Tuning, pues este es superior a sus contrincantes en todas las métricas si a media nos referimos. Ahora bien, si hacemos énfasis en la desviación estándar de cada modelo es posible afirmar que MobileNetV3 no es el modelo de la desviación más baja dado que en algunas métricas el modelo ResNet50V2 cuenta con la menor variabilidad, más sin embargo este modelo no logra superar a MobileNetV3 por que lo que se puede concluir que MobileNetV3 es el modelo que mejor se ajusta al problema tratado.

A continuación, en la tabla 3 se presenta la información total de los mejores resultados de MobileNetV3.

Tabla 3. Mejores resultados de MobileNetV3.

Mejores resultados	
Accuracy	0,9380
Recall	0,9572
Specificity	0,9213
AUC	0,9740
F1 Score	0,9347
MCC	0,8766
Seed	3593322475
Epoch	18

Con el fin de ampliar aún más la capacidad del modelo ganador se procede a realizar un proceso de aumento de datos con variaciones radiométricas al dataset, para esto se generan cinco variaciones de cada imagen con rotaciones ± 20 grados, zoom de $\pm 20\%$, corrimiento horizontales y verticales de $\pm 10\%$ y reflejado de imagen. Posterior a ello se procede a reentrenar el modelo

MobileNetV3 según el plan de pruebas y finalmente en la tabla 4 se presentan los mejores resultados obtenidos en esta ocasión, donde se logra incluso mejorar los resultados obtenidos anteriormente, todo gracias a que este aumento de datos logra nutrir el dataset de mayor variabilidad y durante el entrenamiento del modelo este logra adquirir mejor capacidad para determinar la clasificación acertada ante casos de alta variabilidad.

Tabla 4. Mejores resultados de MobileNetV3 con aumento de datos

Mejores resultados	
Accuracy	0,9433
Recall	0,9403
Specificity	0,9460
AUC	0.9800
F1 Score	0.9390
MCC	0.8861
Seed	3593322475
Epoch	33

Referencias

- [1] ICONTEC. (2009). NORMA TÉCNICA GTC COLOMBIANA 24
- [2] Acoplasticos. (2019). Residuos aprovechables y no aprovechables. [En línea], Disponible: <https://www.acoplasticos.org/boletines/2019/NormasAmbientales/IG-11.%20Residuos%20aprovechables%20y%20no%20aprovechables.pdf>
- [3] Arquitectura, E., Diseño, I. Y., Mena, A. S. UNIVERSIDAD EUROPEA DE MADRID. Clasificación de logos de marcas mediante Deep Learning. <https://titula.universidadeuropea.com/bitstream/handle/20.500.12880/779/SantosMenaAdria.n.pdf?sequence=1>
- [4] Brownlee, Jason. (2019, Diciembre 20). A Gentle Introduction to Transfer Learning for Deep Learning. [En línea] Disponible: <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- [5] Chicco, Davide Jurman, Giuseppe. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics.
- [6] Arquitectura, E., Diseño, I. Y., Mena, A. S. UNIVERSIDAD EUROPEA DE MADRID. Clasificación de logos de marcas mediante Deep Learning.
- [7] Google developers. (2020, February 10). Clasificación: Curva ROC y AUC. [En línea], Disponible: <https://developers.google.com/machine-learning/crash-course/classification/roc-andauc?hl=es419roc-curve>
- [8] Huang-Fu, C. Y., Liao, C. H., Wu, J. Y. (2021). Comparing the performance of machine learning and deep learning algorithms classifying messages in Facebook learning group. Proceedings - IEEE 21st International Conference on Advanced Learning Technologies, ICALT 2021, 347–349. <https://doi.org/10.1109/ICALT52272.2021.00111>