



Clasificador de latas sin barniz interno dentro del proceso de fabricación de envases de aluminio para bebidas

Can classifier for detection of absence of internal varnish within the manufacturing process of aluminum containers for beverages

Juan Nicolas Soto Ríos¹

RESUMEN

Este artículo propone el desarrollo de un clasificador de latas de aluminio sin barniz interno de una línea de producción, a partir del entrenamiento de un modelo pre entrenado denominado MobileNet V2, usando un conjunto de datos específico. Se recogieron 1766 imágenes de latas de una línea de producción, que incluían envases con barniz y sin barniz interno. Se etiquetaron las imágenes de latas con barniz interno con [OK] y aquellas que contenían latas sin barniz con [W]. Se realizó la partición aleatoria de los datos en dos grupos y a partir de un proceso de validación, se aseguró la proporcionalidad de imágenes con etiquetas [W] y [OK] para ambos grupos del DataSet. Posteriormente, se efectuó la preparación de los datos, convirtiendo las imágenes a tensores, donde se cambia el tamaño de las mismas al requerido por el modelo (224x224). A partir del entrenamiento del algoritmo, se consiguió una precisión general de 0,88 y una proporción de falsos negativos para la categoría de latas sin barniz [W] de 0,1. Por tanto, el modelo logra clasificar correctamente el 88% de las imágenes, identificando acertadamente las que tienen barniz y las que no lo tienen. De esta manera, a partir de la aplicación de una herramienta de clasificación de imágenes, impulsada por algoritmos de Deep Learning, se logra dar solución a problemas del área industrial. En este caso específico, la ausencia de barniz

¹ Fundación Universitaria Los Libertadores, Bogotá-Colombia, Contacto: jnsotor@libertadores.edu.co



interno en latas de aluminio, disminuyendo fallas en el proceso y asegurando la calidad del producto final.

Palabras clave: Inteligencia Artificial, Machine Learning, Aprendizaje Profundo, Redes Neuronales Convolucionales, Reconocimiento de Imágenes.

ABSTRACT

This article proposes the development of an aluminum can classifier without internal varnish of a production line, from the training of an existing model called MobileNet V2, using a specific data set. 1766 images of cans were collected from a production line, which included containers with varnish and without internal varnish. Images of cans with internal varnish were labeled with [OK] and those containing cans without varnish with [W]. The data was randomly partitioned into two groups and from a validation process, the proportionality of images with labels [W] and [OK] was ensured for both groups of the DataSet. Subsequently, the data preparation was carried out, converting the images to tensors, where the size of the images is changed to that required by the model (224x224). From algorithm training, an overall precision of 0.88 and a false negative ratio for the unvarnished cans category [W] of 0.1 were achieved. Therefore, the model manages to correctly classify 88% of the images, correctly identifying those that have varnish and those that do not. In this way, from the application of an image classification tool, driven by Deep Learning algorithms, it is possible to solve problems in the industrial area. In this specific case, the absence of internal varnish in aluminum cans, reducing failures in the process and ensuring the quality of the final product.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Networks, Image recognition.

INTRODUCCIÓN

El presente trabajo de investigación pretende detectar la ausencia de aplicación de barniz interno en latas de aluminio para bebidas. La falta de este barniz puede generar cambios en



el sabor de la bebida por reacciones químicas con el aluminio, causando que la experiencia del consumidor final no sea satisfactoria. Esta falla de proceso, que afecta la calidad del producto, puede traer problemas de diferente índole. Con el cliente principal, la embotelladora, pueden darse penalizaciones económicas. Con el cliente final o consumidor, podría traer problemas aún más graves: la afectación de la imagen de la embotelladora y la pérdida del contrato parcial o total, para la planta productora.

De esta manera, el reciente desarrollo de los métodos de análisis inteligente de datos ha traído nuevas oportunidades y retos para una amplia variedad de problemas científicos (Boubacar, Kamsu-Foguem & Tangara, 2018), incluyendo su aplicación en casos como la industria de la manufactura. Por ejemplo, en el aseguramiento de la calidad de los productos. De este modo, una de las herramientas que puede brindar ayuda en esta área son las técnicas de Visión Artificial, por ejemplo, el reconocimiento de objetos que tiene como objetivo determinar si hay alguna instancia de objetos de una categoría dada, en una imagen (Liu, Ouyang, Wang, Fieguth, Liu & Pietikäinen, 2019).

Específicamente, para este trabajo se propone usar una de las herramientas de reconocimiento de objetos y clasificación de imágenes, impulsadas por algoritmos de Deep Learning, como son las Redes Neuronales Convolucionales (CNN). Este tipo de redes están conformadas por múltiples capas internas, que tienen una arquitectura especialmente diseñada para el procesamiento de imágenes (Grinblat, Uzal, Larese & Granitto, 2016). Teniendo en cuenta lo anterior, el objetivo de esta investigación es el diseño de un modelo de reconocimiento de imágenes, que permita detectar piezas fuera de especificaciones técnicas.

Particularmente, este modelo de reconocimiento de imágenes tiene el objetivo de detectar las latas sin barniz en una línea de producción. Específicamente, el modelo logra identificar latas sin aplicación de barniz en un 90% de los casos. Lo anterior reduciría el riesgo de entregar un producto por fuera de las especificaciones y evitar así, los impactos negativos mencionados con anterioridad, tanto para la planta productora, como para el cliente principal.



REFERENTES TEÓRICOS

Inteligencia Artificial

La Inteligencia Artificial se puede definir como una rama de la ciencia de la computación, que incluye el estudio de la relación entre computación y cognición (Barr & Feigenbaum, 1984). De esta manera, la Inteligencia Artificial se ha propuesto dentro de sus objetivos, mimetizar o imitar la inteligencia o patrones de comportamiento de los seres humanos o de cualquier otra entidad viviente. Así pues, esta área de estudio se ha ocupado del desarrollo de computadoras capaces de participar en procesos de pensamiento similares a los humanos, como el aprendizaje, el razonamiento y la autocorrección (Kok, Boers, Kusters, Van der Putten & Poel, 2009).

La investigación que condujo a la tecnología actual de la Inteligencia Artificial (AI) comenzó después de la segunda guerra mundial y fue estimulada en gran parte por la llegada de la computadora digital (Simmons & Chapell, 1988). No obstante, fue hasta el año de 1950 cuando se acuñó el término Inteligencia Artificial para describir la investigación relacionada con la mimetización de comportamientos humanos, a partir del procesamiento de imágenes (Simmons & Chapell, 1988). Después de su aparición, la AI ha transicionado de la teoría a aplicaciones tangibles, como, por ejemplo, el análisis de largas bases de datos en tiempo cercano, autos de conducción autónoma, recomendaciones de compra online, entre otras (Helm, Swiergosz, Haeberle, Karnuta, Schaffer, Krebs, Spitzer & Ramkumar, 2020).

Aprendizaje de Máquina

Dentro de la Inteligencia Artificial, se encuentra el subcampo del Aprendizaje de Máquina, o en inglés, Machine Learning. Esta área tiene el objetivo de construir o mapear una función, a partir del entrenamiento de un conjunto de datos de entrada y salida, que permita realizar predicciones con datos futuros (Wang, 2016). En otras palabras, Machine Learning construye algoritmos computacionales a partir del input y output de datos conocidos, que le



permite aprender, mejorar sus análisis y de esta manera, crear recomendaciones o tomar decisiones (Bini, 2018).

Si bien Machine Learning ha contribuido al desarrollo de la sociedad moderna con su utilidad en áreas como el comercio online, páginas web, redes sociales, entre otros campos de aplicación, sus técnicas convencionales presentan limitaciones importantes. Con la aparición de datos de entrenamiento a gran escala, el rendimiento de modelos de Machine Learning con estructuras superficiales se satura, por su capacidad de aprendizaje limitada (Wang, 2014). Adicionalmente, la ingeniería convencional de Machine Learning presenta dificultades para procesar conjuntos de datos naturales en su forma cruda. En consecuencia, el Aprendizaje Profundo resulta ser una técnica que requiere poca ingeniería a mano, por lo que puede aprovechar el aumento en la cantidad de datos y cálculos disponibles (LeCun, Bengio & Hinton, 2015).

Deep Learning

Deep Learning es una forma de Machine Learning que permite a las computadoras aprender de la experiencia y entender el mundo en términos de una jerarquía de conceptos (Goodfellow, 2016). Específicamente, el Aprendizaje Profundo es un conjunto de algoritmos que intentan aprender en múltiples niveles, que corresponden a diferentes categorías de abstracción (Deng & Yu, 2014). Los algoritmos de Deep Learning se inspiran y están diseñados de tal forma, que imitan la función de la corteza cerebral humana (Chauhan, Ghansla & Joshi, 2018). Es decir, que están inspirados en la estructura y función de las redes neuronales cerebrales.

Redes Neuronales Convolucionales

Teniendo en cuenta lo anterior, el Aprendizaje Profundo o Deep Learning se basa en el uso de Redes Neuronales Artificiales. Particularmente, en el empleo de Redes Neuronales Convolucionales. Este tipo de redes constituyen un sistema de neuronas que intercambian mensajes entre sí (Yi, 2018). Fundamentalmente, este tipo de redes están compuestas por múltiples capas: Capas convolucionales, que extraen diferentes características de los datos de entrada; Capas de agrupación, que extraen características libres de ruido o distorsión, y



capas completamente conectadas, que proyectan las características a una capa clasificadora para etiquetar el dato de entrada o retroalimentar, para mejorar los parámetros (Yi, 2018).

Las redes Neuronales Convolucionales, tienen una amplia aplicación en diversos dominios como, por ejemplo, análisis de señales auditivas y visuales (Russakovsky et al., 2015), reconocimiento de desastres (Liu and Wu, 2016), reconocimiento de voces, visión por computador (Karpathy et al., 2014) y procesamiento de lenguaje automatizado (Hinton et al., 2006). Especialmente, en el área de la visión computacional y el reconocimiento de objetos, que pretende localizar y clasificar objetos existentes en cualquier imagen y etiquetarla (Zhao, Zheng, Xu & Wu, 2019), tiene gran utilidad. Así pues, este tipo de redes muestran grandes ventajas cuando tratan con grandes regiones de imágenes, con variaciones complejas, porque su arquitectura profunda puede aprender mejor la representación de características contextuales globales (Wang, 2016).

METODOLOGÍA

Para el desarrollo de esta investigación se utilizó una metodología estadística como se muestra a continuación.

Construcción de la base de datos

Para poder entrenar el modelo de clasificación de imágenes, se recogieron 1766 imágenes en las líneas de producción. Estas imágenes fueron tomadas en un espacio con condiciones controladas, con el fin de captar suficientes imágenes, donde se pudieran diferenciar las dos categorías a clasificar, con barniz y sin barniz.

Figura 1 Fuente: Elaboración propia.



Figura 2 Fuente: Elaboración propia.





LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

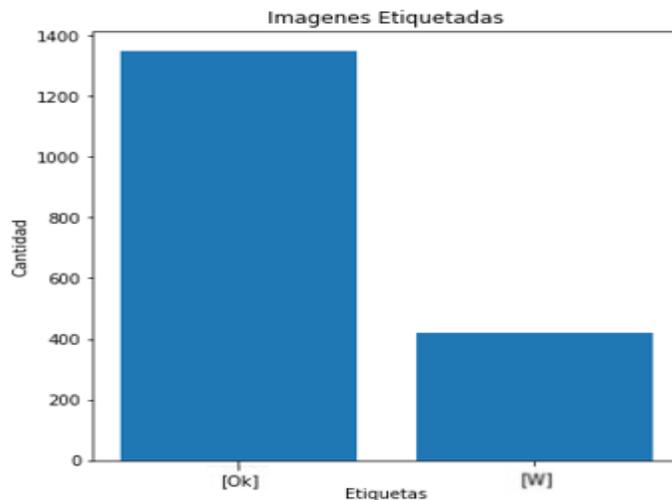


Etiquetado de Imágenes

Se etiquetaron con [OK] las latas que tienen barniz interno y con [W] las imágenes donde se tuvieran una o más latas sin aplicación de barniz, con una proporción de categorías como se muestra a continuación.



Figura 3 Frecuencia de etiquetas. Fuente: Elaboración propia.



División del conjunto de datos

Se realizó una partición aleatoria de los datos recolectados en dos grupos; el primero, correspondiente a los datos de entrenamiento que contienen el 80% de las imágenes y el segundo, que contiene el 20% de imágenes restantes, correspondiente a los datos de validación. Con este último grupo, se realizó la evaluación del modelo para obtener las métricas de desempeño.

A partir de la partición del conjunto de datos, se validó que tanto el grupo de entrenamiento, como el grupo de validación, mantuvieran la proporción de etiquetas [OK] y [W] como se muestra a continuación: Proporción de datos [W] de entrenamiento: 31.10% y proporción de [W] de validación: 31.11%.

Descripción del modelo

Se escogió un modelo preentrenado del repositorio de Tensor Flow llamado MobileNet V2, este modelo pertenece a una familia de arquitectura de redes neuronales para tareas de



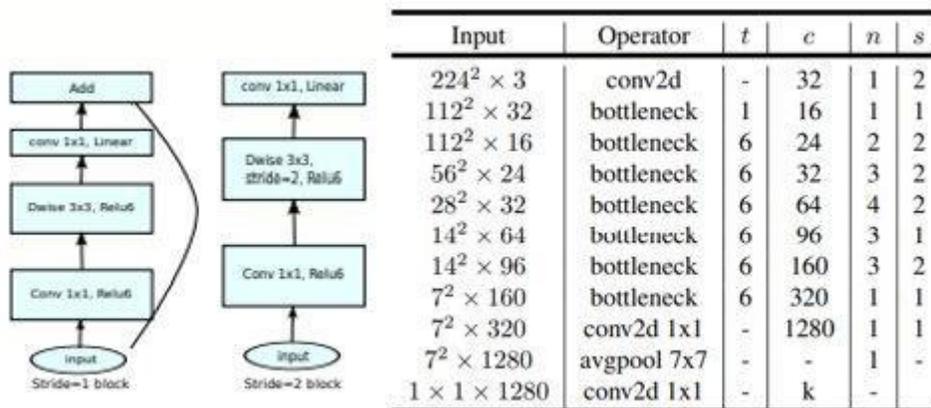
clasificación de imágenes, publicado por Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, Liang-Chieh Chen (2018).

Posteriormente se hizo el *fine tuning* del modelo, es decir, la personalización del modelo preentrenado conectando la última capa del modelo con una capa de salida de tamaño 2, correspondiente al número de etiquetas, con una función de salida sigmoide, optimizador “Adam” y adicionalmente se agregó un callback que detenga el modelo cuando el Accuracy sobre los datos de validación se estabilice durante tres iteraciones del modelo, esto último con el fin de evitar el overfitting del modelo sobre los datos de entrenamiento.

Para este modelo se va a utilizar una capa de entrada de $224 \times 224 \times 3$, tamaño al que habrá que llevar las imágenes con las que se va a alimentar el modelo y una capa de salida de tamaño 2, que corresponde a las etiquetas de las imágenes [Ok] y [W].

A continuación, se muestra la arquitectura del modelo.

Figura 4 Arquitectura del modelo MobilNet V2 (Roslidar, Saddami, et al., 2019)



Preprocesamiento de los datos



Una vez definido y compilado el modelo, se procede a preprocesar el set de datos de entrenamiento. Esta etapa corresponde a convertir las imágenes a Tensores, que son una representación vectorial de las imágenes y cambiar el tamaño de las mismas, al requerido por el modelo (224x224).

RESULTADOS

Entrenamiento del modelo

Se entrena el modelo con los datos de entrenamiento. A continuación, se muestra la evolución de accuracy y error.

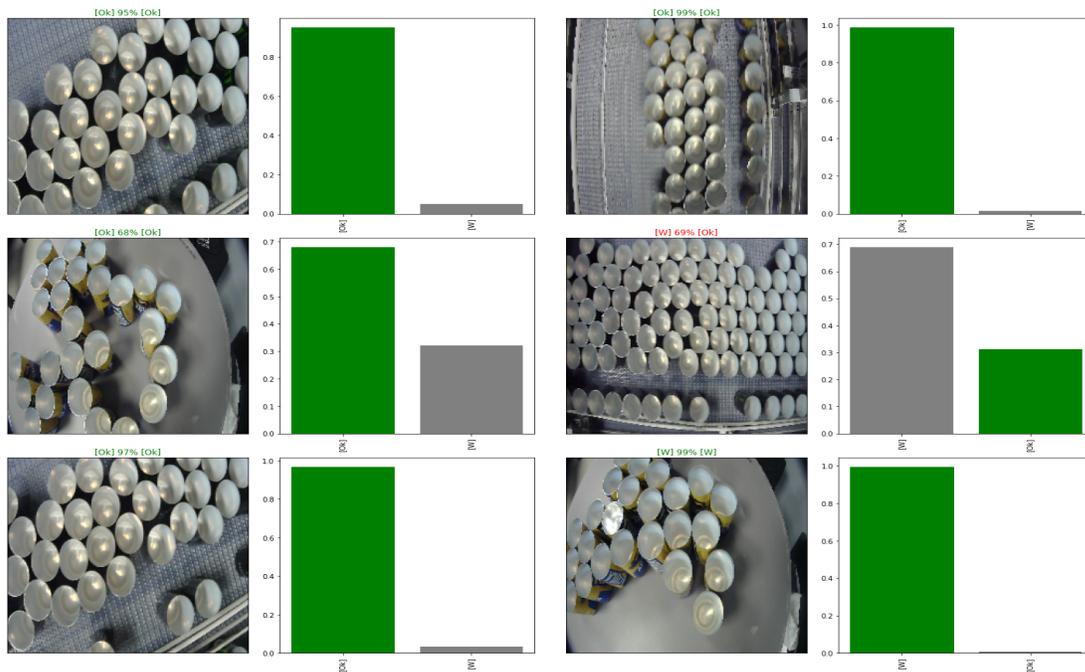
Figura 5 Evolución de métricas del modelo. Fuente: Elaboración propia.





Una vez entrenado el modelo se hicieron varias predicciones, validando la etiqueta real y las probabilidades de cada una de ellas.

Figura 6 Predicciones del modelo sobre el set de validación. Fuente: Elaboración propia.

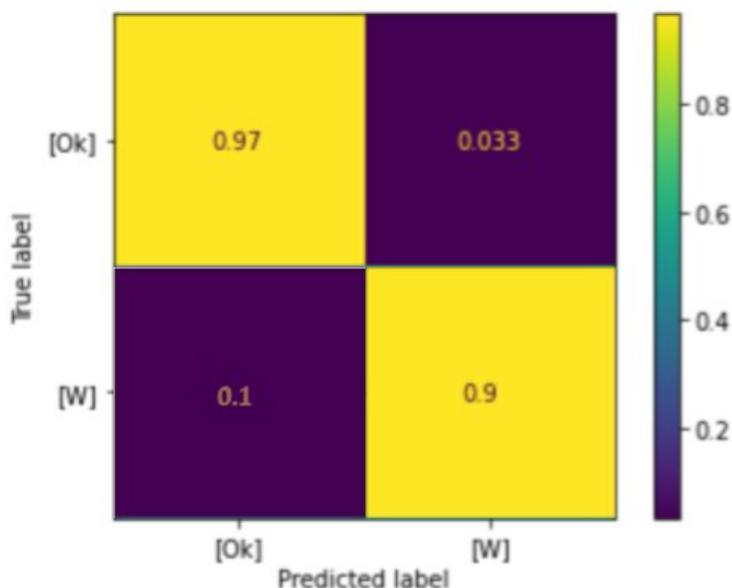


1. Desempeño del modelo

A continuación, se muestra la matriz de confusión donde se comparan las etiquetas reales con las predichas por el modelo, todo esto sobre el set de validación.



Figura 7 Matriz de confusión. Fuente: Elaboración propia.



De la de la matriz de confusión se puede observar la tasa de falsos positivos de 0,1 arrojados por el modelo, es decir, el modelo identifica un 90% de las latas sin barniz que se le muestran.

Así mismo, se corrió el reporte de desempeño obteniendo los siguientes resultados:

Figura 8 Reporte de desempeño. Fuente: Elaboración propia.

	precision	recall	f1-score	support
[Ok]	0.97	0.87	0.92	270
[W]	0.69	0.90	0.78	84
accuracy			0.88	354
macro avg	0.83	0.89	0.85	354
weighted avg	0.90	0.88	0.89	354

2. Conclusiones

En este artículo se abordó el desarrollo de un clasificador de latas para bebidas sin barniz interno, de una línea de producción. Primero se creó un conjunto de datos a partir de la toma de imágenes de latas, que incluían envases con barniz y sin barniz. Segundo, se llevó



a cabo la etiquetación de las imágenes con dos rótulos: [OK] para latas con barniz y [W] para imágenes que contenían envases sin barniz interno. En tercer lugar, se elaboró la partición del conjunto de datos, incluyendo el aseguramiento de la proporcionalidad de imágenes con las etiquetas anteriores, para ambos grupos del DataSet. Cuarto, se preprocesaron los datos y finalmente se entrenó el modelo.

A partir del modelo desplegado se consiguió una precisión general de 0,88. Por tanto, el modelo logra clasificar correctamente el 88% de las imágenes, identificando acertadamente las que tienen barniz y las que no tienen barniz. Adicionalmente, la proporción de falsos negativos obtenidos para la categoría de latas sin barniz [W] fue de 0,1 cumpliendo con el mínimo esperado. Lo anterior es importante, ya que el clasificador debe ser capaz de detectar la mayor cantidad de latas sin barniz, porque un falso positivo, es decir, clasificar la lata como [Ok] cuando debería ser [W] es un evento que pondría en riesgo la calidad del producto.

Finalmente, se recomienda para futuros usos de este modelo y su posterior desarrollo, utilizar algoritmos de *Data Augmentation* o aumentar el tamaño del DataSet. Lo anterior, con el fin de lograr mejores métricas del modelo.

REFERENCIAS BIBLIOGRÁFICAS

- Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. *The Journal of arthroplasty*, 33(8), 2358-2361.
- Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018, December). Convolutional neural network (CNN) for image detection and recognition. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 278-282). IEEE.



- Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and trends in signal processing*, 7(3-4), 197-387.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
- Grinblat, G. L., Uzal, L. C., Larese, M. G., & Granitto, P. M. (2016). Deep learning for plant identification using vein morphological patterns. *Computers and Electronics in Agriculture*, 127, 418-424.
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., ... & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: Definitions, applications, and future directions. *Current reviews in musculoskeletal medicine*, 13(1), 69-76.
- Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).
- Kok, J. N., Boers, E. J., Kusters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence*, 1, 270-299.
- Liu, X., Liu, W., Ma, H., & Fu, H. (2016, July). Large-scale vehicle re-identification in urban surveillance videos. In *2016 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 1-6). IEEE.
- Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep learning for generic object detection: A survey. *International journal of computer vision*, 128(2), 261-318.



- Roslidar, R., Saddami, K., Arnia, F., Syukri, M., & Munadi, K. (2019, August). A study of fine-tuning CNN models based on thermal imaging for breast cancer classification. In *2019 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)* (pp. 77-81). IEEE.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), 211-252.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510-4520).
- Simmons, A. B., & Chappell, S. G. (1988). Artificial intelligence-definition and practice. *IEEE Journal of Oceanic Engineering*, *13*(2), 14-42.
- Traore, B. B., Kamsu-Foguem, B., & Tangara, F. (2018). Deep convolution neural network for image recognition. *Ecological Informatics*, *48*, 257-268.
- Wang, X. (2016). Deep learning in object recognition, detection, and segmentation. *Foundations and Trends in Signal Processing*, *8*(4), 217-382.
- Yi, Z. (2018, September). Evaluation and implementation of convolutional neural networks in image recognition. In *Journal of Physics: Conference Series* (Vol. 1087, No. 6, p. 062018). IOP Publishing.
- Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, *30*(11), 3212-3232.