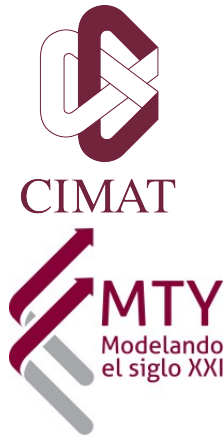


Aumento Automático de DATOS



Nelson Ariza Morales
CIMAT
Unidad Monterrey

El presente trabajo se centra en la implementación y evaluación de técnicas de aumento automático de datos para mejorar el rendimiento de modelos de clasificación de imágenes. Se utilizó el modelo ResNet-18, en conjunto con herramientas avanzadas como AutoAugment y Optuna. Estas herramientas permitieron optimizar de manera automática las políticas de transformación aplicadas a los datos del dataset CIFAR-10. El aumento automático de datos incrementó la diversidad del conjunto de entrenamiento mediante transformaciones como rotaciones, cambios de brillo y cizalladuras, entre otras. Además, el uso de Optuna facilitó la selección de los hiperparámetros óptimos para estas transformaciones.

Introducción

EN los campos de la visión por computadora y el aprendizaje automático, la disponibilidad y calidad de los datos representan un desafío constante. Problemas como la falta de datos, el desbalance entre clases, la presencia de ruido, y las restricciones logísticas y económicas en la recolección de información limitan la capacidad de los modelos para generalizar y resolver tareas complejas de manera efectiva. Estos obstáculos son especialmente críticos en escenarios donde los datos disponibles no reflejan adecuadamente las variaciones que pueden encontrarse en aplicaciones del mundo real.

El tratamiento adecuado de los datos es un paso esencial en cualquier pipeline de aprendizaje automático, ya que impacta directamente en la capacidad del modelo para aprender patrones significativos. Antes del modelado, es crucial abordar problemas como la limpieza de datos, la selección de características relevantes y, en particular, el aumento de la diversidad del conjunto de datos mediante técnicas como el aumento automático de datos. Estas estrategias permiten enriquecer el conjunto de entrenamiento con variaciones artificiales que simulan condiciones reales, mejorando así la robustez y generalización del modelo.

En este contexto, AutoML (Automated Machine Learning) ha emergido como un enfoque poderoso para automatizar las decisiones críticas en el proceso de aprendizaje, incluida la selección de transformaciones para el aumento de datos. Herramientas como AutoAugment y algoritmos de optimización como Optuna, AutoPytorch permiten explorar y seleccionar automáticamente las mejores políticas de transformación, reduciendo la necesidad de intervención manual y asegurando configuraciones óptimas.

Este trabajo se enfoca en implementar técnicas de aumento automático de datos en el dataset CIFAR-10, utilizando el modelo ResNet-18 como base para la clasificación de imágenes. Mediante el uso de AutoAugment y la optimización de políticas con Optuna, se logró incrementar la diversidad del conjunto de entrenamiento, optimizar hiperparámetros y evaluar el impacto de estas transformaciones en el rendimiento del modelo. Los resultados obtenidos no solo demuestran la efectividad del aumento automático, sino que también subrayan la importancia de integrar enfoques de AutoML para mejorar la generalización y robustez de los modelos en problemas de clasificación de imágenes.

Objetivo

Desarrollar un aumento automático de datos con técnicas de AutoML para mitigar el desbalanceo en el dataset CIFAR-10 y mejorar la generalización del modelo de clasificación.

Marco teórico

En aprendizaje automático, uno de los desafíos más comunes es la insuficiencia o desbalance en los datos disponibles, lo cual limita la capacidad de los modelos para generalizar y enfrentar variaciones reales. Para abordar este problema, el aumento automático de datos surge como una técnica que aplica transformaciones sistemáticas a los datos originales, incrementando así su diversidad. Herramientas como *AutoAugment* y *RandAugment* han demostrado su eficacia al optimizar políticas de transformación que no solo mejoran el rendimiento del modelo, sino que también reducen la necesidad de intervención manual.

Las imágenes, al ser procesadas por computadoras, se representan como matrices numéricas. Cada píxel corresponde a un valor que indica su intensidad en uno o más canales:

- En imágenes a escala de grises, cada píxel es un valor escalar.
- En imágenes a color (RGB), cada píxel se describe mediante tres valores, correspondientes a los canales rojo, verde y azul.

Transformaciones en la Representación de Imágenes



Figura 1: Aumentar la nitidez

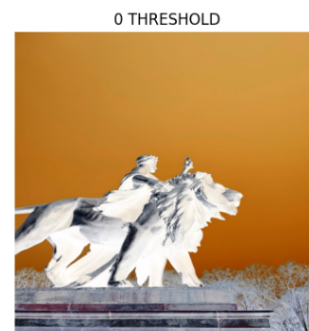


Figura 2: Invertir los pixeles



Figura 3: Rotar



Figura 4: Aumentar brillo

Las transformaciones aplicadas durante el aumento de datos afectan directamente las matrices que representan las imágenes:

- **Geométricas:** Una rotación reorganiza los píxeles, cambiando su posición relativa; la traslación desplaza la imagen dentro del marco definido por la matriz; la cizalladura deforma los ejes, alterando la forma de los objetos representados.
- **Fotométricas:** Un ajuste de brillo escala los valores de intensidad de los píxeles; el contraste modifica la relación entre valores altos y bajos; la inversión cambia los valores de los píxeles al complementar su intensidad.

Estas transformaciones no solo amplían la variedad del dataset, sino que también permiten al modelo aprender características invariantes a cambios comunes en las imágenes.

Optimización de Políticas

El aumento automático de datos no se limita a aplicar transformaciones simples, sino que busca optimizar las combinaciones de transformaciones para maximizar el rendimiento del modelo. Herramientas como *AutoAugment* automatizan este proceso mediante:

- La definición de políticas, que combinan transformaciones con intensidades y probabilidades específicas.
- Algoritmos de búsqueda, como el aprendizaje por refuerzo o la optimización bayesiana, que identifican las políticas más efectivas.
- Evaluaciones iterativas, donde las políticas se prueban en el modelo y se seleccionan las mejores con base en métricas de rendimiento.

Este enfoque asegura que las transformaciones aplicadas sean relevantes para el problema, evitando configuraciones arbitrarias.

ResNet-18 en la Clasificación de Imágenes

ResNet-18 es una red neuronal profunda diseñada para resolver problemas de degradación del gradiente mediante bloques residuales. Estos bloques permiten que las capas aprendan características más complejas sin perder información clave de las capas previas.

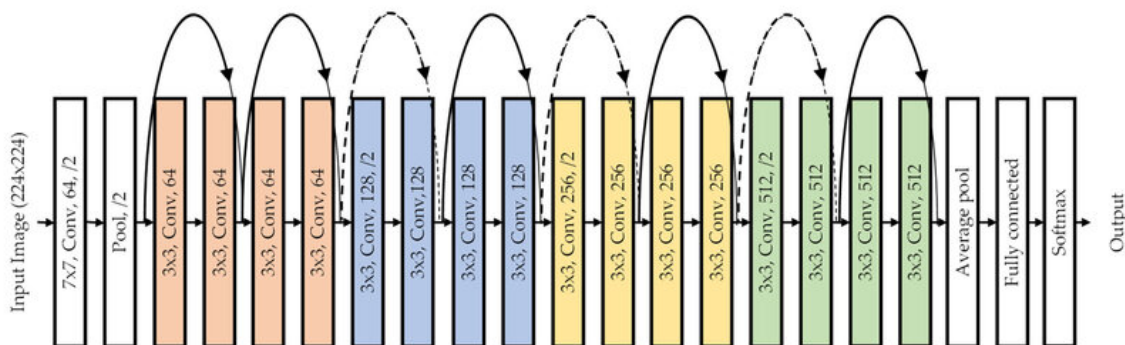


Figura 5: Configuración del modelo ResNet-18

En este proyecto, ResNet-18 se utilizó como modelo base para clasificar las imágenes del dataset CIFAR-10. Para ajustarlo a esta tarea, se reemplazó su capa final completamente conectada,

adaptándola a las 10 clases del dataset. La capacidad de ResNet-18 para extraer características jerárquicas lo convierte en una elección adecuada para evaluar el impacto de las políticas de aumento automático.

Optuna y la Optimización de Hiperparámetros

La automatización del aumento de datos requiere herramientas que permitan explorar configuraciones de manera eficiente. En este proyecto, *Optuna* se utilizó para optimizar hiperparámetros como la intensidad y probabilidad de las transformaciones. A diferencia de enfoques manuales, Optuna aplica técnicas avanzadas, como la optimización bayesiana, para buscar las mejores configuraciones.

Al integrarse en el pipeline, Optuna permitió seleccionar políticas que incrementaron la precisión del modelo, demostrando su capacidad para complementar técnicas de aumento automático y mejorar significativamente el rendimiento en tareas de clasificación de imágenes.

Metodología

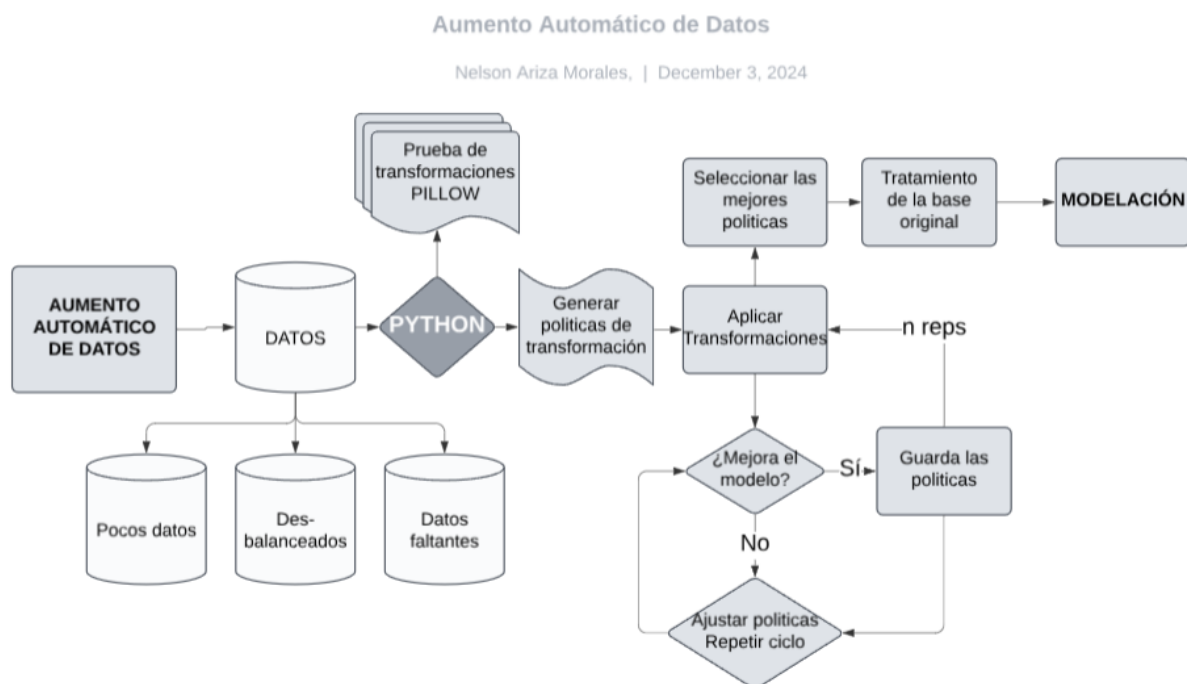


Figura 6: Diagrama de flujo del método

El desarrollo del proyecto siguió un flujo bien definido para implementar y optimizar técnicas de aumento automático de datos. Se utilizó el dataset CIFAR-10, compuesto por imágenes de 32×32 píxeles distribuidas en 10 clases. Este conjunto se dividió en datos de entrenamiento y prueba. Previo al modelado, se aplicaron transformaciones básicas para normalizar las imágenes y mejorar su representatividad.

Se seleccionó la arquitectura *ResNet-18* como modelo base, ajustando su capa final completamente conectada para clasificar las 10 clases del dataset. Este modelo preentrenado en ImageNet sirvió como referencia para evaluar el impacto de las políticas de aumento en su rendimiento. Se definieron políticas iniciales de transformación que combinaban ajustes geométricos (rotación,

cizalladura, traslación) y fotométricos (cambio de brillo, contraste). Estas transformaciones se aplicaron de forma aleatoria para generar un conjunto inicial de políticas.

Para seleccionar las mejores políticas de aumento, se utilizó *Optuna*, una biblioteca de optimización de hiperparámetros. El proceso incluyó:

- **Definición de hiperparámetros:** Se configuraron parámetros como la intensidad de las transformaciones y la probabilidad de aplicación.
- **Evaluación iterativa:** Cada conjunto de políticas generado se probó entrenando el modelo con los datos transformados, evaluando su precisión en el conjunto de validación.
- **Selección de políticas:** Optuna identificó las combinaciones de transformaciones que maximizaron el rendimiento del modelo.

Una vez seleccionadas las políticas óptimas, se aplicaron al conjunto de entrenamiento completo para generar un dataset aumentado. Este nuevo conjunto, enriquecido con variaciones generadas automáticamente, fue utilizado para entrenar nuevamente el modelo ResNet-18, obteniendo un modelo final más robusto.

El modelo entrenado con el dataset aumentado se evaluó en el conjunto de prueba, comparando su precisión con la obtenida sin aumento de datos. Esta comparación permitió cuantificar la mejora derivada de las políticas optimizadas.

El flujo metodológico integró la definición de transformaciones, su optimización mediante hiperparámetros y la evaluación iterativa en el modelo base, asegurando un proceso eficiente y dirigido a mejorar la generalización del modelo.

Resultados y discusiones

Baseline: Modelo con Dataset Desbalanceado

Inicialmente, el dataset CIFAR-10 se desbalanceó reduciendo significativamente la cantidad de muestras en ciertas clases, como *camiones*, *Barcos* y *gatos*, mientras que otras clases se mantuvieron con una representación completa. Con esta base desbalanceada, se entrenó el modelo *ResNet-18* para establecer una línea base (*baseline*) de rendimiento. Cabe mencionar que la configuración base del modelo resnet18 se corrió por 10 épocas debido al recurso computacional y de tiempo con el que se contaba

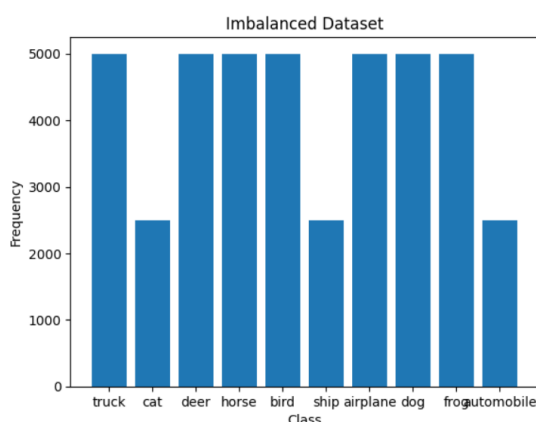


Figura 7: Base desbalanceada

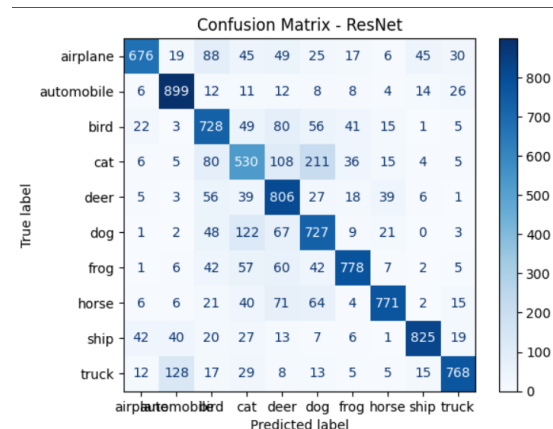


Figura 8: Matriz de confusión de *BaseLine*

La evaluación del modelo mostró una precisión (*accuracy*) de 70.81 %, evidenciando que las clases con menos representaciones, como *carros* y *gatos*, se clasificaban incorrectamente en su mayoría, lo que se reflejó en la matriz de confusión.

Posteriormente, se aplicó el proceso de optimización de políticas de aumento utilizando *Optuna*. Este proceso seleccionó combinaciones óptimas de transformaciones geométricas (rotación, cizalladura, traslación) y fotométricas (brillo, contraste), maximizando la precisión del modelo en cada iteración. Las mejores políticas identificadas se utilizaron para aumentar las clases subrepresentadas del dataset, logrando una base balanceada en términos de distribución de clases. Podemos notar tanto en la figura 9 como en la figura 10 que realmente la optimización para seleccionar las mejores políticas ha resultado fructífera ya que las políticas siguen siendo funcionales para la naturaleza del problema que es la clasificación, ya que en la mayoría de ellas vemos claramente la distinción del objeto sobre la demás parte de la imagen. Entonces, a primera vista notamos que el aumento de datos puede ser beneficioso, pero aún debemos probar con el modelo resnet18 nuevamente para estar seguros de que ha mejorado realmente.

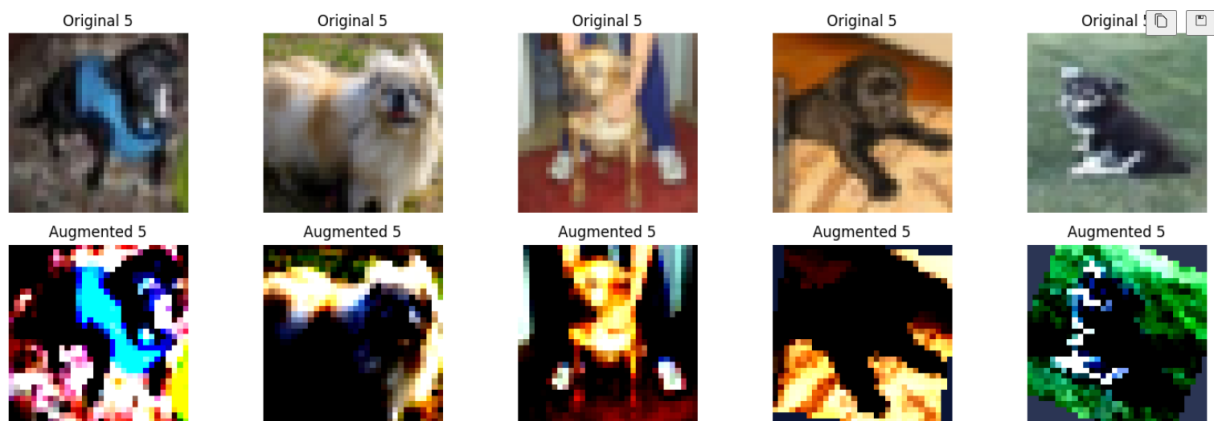


Figura 9: Aumento de datos en la clase de perros

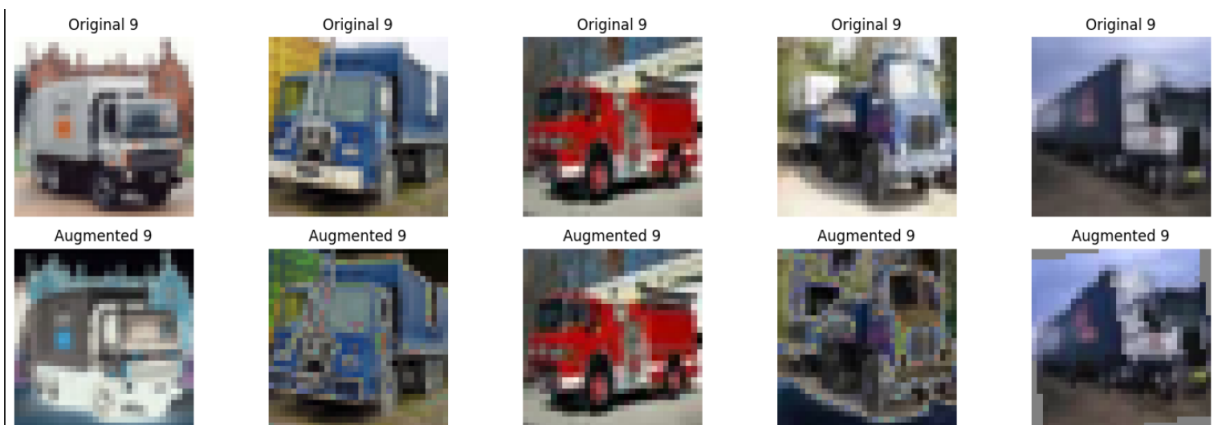


Figura 10: Aumento de datos en la clase de camiones

En la primera fila se muestran ejemplos originales de la clase "perros" del dataset CIFAR-10. En la segunda fila, se presentan las imágenes generadas utilizando las políticas optimizadas de aumento automático. Se observa cómo las transformaciones aplicadas, como rotaciones, cizalladuras y ajustes de brillo, generan variaciones visuales que enriquecen el conjunto de datos, sin alterar la identidad de la clase.

Similarmente, en la primera fila se encuentran imágenes originales de la clase camiones”. Las imágenes en la segunda fila muestran resultados del aumento, donde transformaciones como traslaciones y cambios en contraste permiten diversificar los datos de esta clase.

Las imágenes generadas mantienen la coherencia semántica con las clases originales, un aspecto crucial en el aumento automático de datos. Las políticas seleccionadas son efectivas para introducir diversidad sin afectar la representatividad de las clases.

```
Epoch [1/10], Loss: 0.9999
Epoch [2/10], Loss: 0.6962
Epoch [3/10], Loss: 0.5921
Epoch [4/10], Loss: 0.4448
Epoch [5/10], Loss: 0.3766
Epoch [6/10], Loss: 0.3501
Epoch [7/10], Loss: 0.2624
Epoch [8/10], Loss: 0.2316
Epoch [9/10], Loss: 0.1504
Epoch [10/10], Loss: 0.1556
Accuracy: 79.01%
79.01
```

Con la base balanceada generada automáticamente, se entrenó nuevamente el modelo *ResNet-18*. El modelo mostró una mejora significativa en el rendimiento, alcanzando una precisión (*accuracy*) de 79.01 % nuevamente en el modelo resnet18 con 10 épocas. Esto demuestra la efectividad del aumento automático de datos para abordar problemas de desbalance de clases. El proceso de aumento automático de datos no solo balanceó la base de datos inicial, sino que también mejoró el rendimiento del modelo al incrementar su capacidad para generalizar sobre clases desbalanceadas.

Conclusión y áreas de mejora

El presente trabajo implementó y evaluó técnicas de aumento automático de datos en el dataset CIFAR-10 utilizando herramientas como AutoAugment y Optuna para optimizar políticas de transformación. El proceso permitió abordar el desbalanceo de clases en el dataset, incrementando la diversidad del conjunto de datos y mejorando significativamente la precisión del modelo base ResNet-18, alcanzando un *accuracy* del 79.01 %, lo que representa una mejora del 9 % en comparación con la línea base establecida con datos desbalanceados, una mejora increíble para solo haber balanceado la base. Se considera importante hacer un análisis de cuáles son las transformaciones que están afectando más a la precisión del modelo y con eso hacer nuevamente una selección con posibles algoritmos evolutivos.

Este trabajo buscó ofrecer una visión integral del proceso de aumento automático de datos, desde la optimización de políticas hasta la evaluación de su impacto en el modelo. Si bien actualmente existen herramientas y modelos optimizados que simplifican y automatizan este proceso, la exploración manual y ajustada permitió comprender mejor cómo cada transformación afecta al modelo y contribuye a su capacidad para generalizar.

En conclusión, el aumento automático de datos es una herramienta poderosa para resolver problemas de desbalanceo y generalización en visión por computadora. Este enfoque no solo mejora el rendimiento de los modelos, sino que también abre la puerta a futuras investigaciones en la automatización y optimización de estos procesos para problemas más complejos.