



**I
N
A
O
E**

**CLASIFICACIÓN DE CASCADAS EXTENSAS DE AIRE DE
ALTAS ENERGÍAS MEDIANTE REDES NEURONALES
CONVOLUCIONALES**

Línea de investigación: Observación de la Tierra

Por

Xochitl Verónica Silvestre Gutiérrez

Propuesta de tesis sometida como requisito parcial para obtener el
grado de

**MAESTRA EN CIENCIAS EN LA ESPECIALIDAD DE
CIENCIA Y TECNOLOGÍA DEL ESPACIO**

En el

Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE)
Santa María Tonantzintla, CP 72840, Puebla, México

Supervisada por:

Dra. Raquel Díaz Hernández, INAOE

Diciembre, 2024

©INAOE 2024

All rights reserved.





“Si deseamos el progreso armónico de nuestra patria debemos estimular tanto la investigación aplicada a fines inmediatos como la investigación en las ciencias fundamentales.”

Manuel Sandoval Vallarta

(Ciudad de México, 11 de febrero de 1899 - 18 de abril de 1977)



AGRADECIMIENTOS

Quiero agradecer a Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por financiar la beca para realizar esta tesis de maestría.

Estoy agradecida con mi asesora la Dra. Raquel Diaz Hernández y al Dr. Leopoldo Altamirano por su constante apoyo con todas las consultas que surgieron y orientar el trabajo por el rumbo correcto.

A mis compañeros del instituto, Mauricio, Andrés y Alan, gracias por su constante ánimo, palabras de aliento y camaradería en los momentos más retadores de esta etapa. Su presencia no solo hizo más llevadero el arduo camino de la maestría, sino que también se convirtió en una fuente de inspiración que me empujó a seguir adelante y no rendirme.

Finalmente, pero no menos importante, mi más profundo reconocimiento y agradecimiento van dirigidos a mi familia: mis padres, hermanos y sobrinos, a Luis Octavio e igualmente a mis mascotas Luna Ximena y Charlotte Carlota. Su amor incondicional, su apoyo constante y su fe en mis capacidades me acompañaron en cada paso de este trayecto.

ABSTRACT

This work presents a methodology for classifying extensive air showers generated by cosmic rays. The proposed approach focuses on using deep convolutional neural networks to extract essential information about primary and secondary particles, thereby improving the accuracy of particle identification.

The research is situated within the context of studying cosmic rays and high-energy cascades, a field that offers a unique perspective for exploring extreme physical phenomena in the universe. However, one of the primary challenges lies in the accurate interpretation of the obtained data, which is hindered by background noise, such as that caused by hadrons. To address this challenge, synthetic data generated through simulations in CORSIKA 7.7550 is utilized, enabling the recreation of high-energy cosmic ray cascades and the generation of images in XZ, YZ, and XY views. These images form a robust database for developing and training the proposed classification methodology.

The neural network-based approach aims to optimize the separation between gamma rays and hadrons within the cascades, therefore facilitating the identification of primary particles. This advancement significantly impacts high-energy astrophysics, opening new opportunities for the rapid classification of such phenomena.

Experiments were conducted with three neural network architectures (ResNet-50, Compact, and AlexNet), evaluating their performance in classifying primary particles (iron, proton, positive muon, and gamma) under different energy levels and zenith angles. The images were analyzed in both color and grayscale, with individual and general views in the XZ, YZ, and XY planes. AlexNet demonstrated the best accuracy and inference speed, achieving 100% accuracy in some color configurations while maintaining an optimal balance between precision and processing time, with faster inference times than ResNet-50. Although ResNet-50 achieved comparable accuracy in some views, its inference times were significantly longer, making it less suitable for applications requiring time efficiency. Compact exhibited good inference times and competitive accuracy, though slightly lower in metrics such as precision and F1-Score, particularly in grayscale configurations.



Overall, the network effectively identified particles such as iron and positive muons, though it faced greater challenges in classifying protons and gamma rays due to the complexity of their cascade morphology and the processes of decay and interaction with the atmosphere.

Keywords: Convolutional networks, High energies, shower classification

RESUMEN

Este trabajo presenta una metodología para la clasificación de cascadas extensas de aire generadas por rayos cósmicos. La propuesta se centra en el uso de redes neuronales convolucionales profundas como herramientas para extraer información esencial sobre partículas primarias y secundarias, mejorando la precisión en la identificación de estas partículas.

La investigación se enmarca en el contexto del estudio de rayos cósmicos y cascadas de altas energías, un campo que ofrece una perspectiva única para explorar fenómenos físicos extremos en el universo. No obstante, uno de los principales retos es la correcta interpretación de los datos obtenidos, que se ve dificultada por el ruido de fondo, como el causado por los hadrones. Para abordar este desafío, se emplean datos sintéticos generados mediante la simulación en CORSIKA 7.7550, que permiten recrear cascadas de rayos cósmicos de altas energías y obtener imágenes en vistas XZ, YZ y XY. Estas imágenes constituyen una base de datos robusta para el desarrollo y entrenamiento de la metodología de clasificación propuesta.

El enfoque basado en redes neuronales busca optimizar la separación entre rayos gamma y hadrones dentro de las cascadas, lo cual facilitaría una identificación de las partículas primarias. Este avance tiene un impacto significativo en el ámbito de la astrofísica de altas energías, abriendo nuevas oportunidades para la rápida clasificación de estos fenómenos.

Se realizaron pruebas con tres arquitecturas de red (ResNet-50, Compact y AlexNet), evaluando su rendimiento en la clasificación de partículas primarias (hierro, protón, muón positivo y gamma) bajo diferentes energías y ángulos cenitales. Las imágenes se analizaron en color y en escala de grises, con vistas individuales y generales en los planos XZ, YZ y XY. AlexNet demostró el mejor desempeño en cuanto a precisión y velocidad de inferencia, alcanzando el 100% de precisión en algunas configuraciones a color y manteniendo un balance óptimo entre precisión y tiempo de procesamiento, con inferencias más rápidas que ResNet-50. Aunque ResNet-50 alcanzó una precisión cercana en algunas vistas, sus tiempos



de inferencia fueron considerablemente mayores, lo cual la hace menos adecuada para aplicaciones donde se requiere reducción de tiempo. Compact mostró buenos tiempos de inferencia y una precisión competitiva, aunque ligeramente inferior en métricas como precisión y F1-Score, especialmente en configuraciones en escala de grises.

En términos generales, la red logró identificar con facilidad partículas como hierro y muones positivos, aunque presentó mayores desafíos en la clasificación de protones y rayos gamma, debido a la complejidad en la morfología de sus cascadas y los procesos de decaimiento e interacción con la atmósfera.

Palabras clave: Redes convolucionales, Altas energías, Clasificación de cascadas.



TABLA DE CONTENIDO

AGRADECIMIENTOS	3
ABSTRACT	4
RESUMEN	6
TABLA DE CONTENIDO	8
ÍNDICE DE TABLAS	11
ÍNDICE DE FIGURAS	12
CAPÍTULO I. INTRODUCCIÓN	13
1.1 Motivación	13
1.2 Justificación	13
1.3 Objetivos	14
1.3.1 Objetivo general	14
1.3.2 Objetivos específicos	14
1.4 Estructura de tesis	15
CAPÍTULO II. ESTADO DEL ARTE	17
2.1 Clasificación con Random Forest	17

2.2 Compactación por Pares.....	18
2.2.1 Adaptación de dominio	19
2.2.2 Clasificación en HAWC	20
CAPÍTULO III. MARCO TEÓRICO	23
3.1 CORSIKA 7.7550	23
3.2 Interacción de las partículas	24
3.3 Cascadas Atmosféricas de Aire.....	25
3.4 Rayos gamma (rayos-γ) como partículas primarias	26
3.4.1 MVTec Deep Learning Tool 24.05.1	27
CAPÍTULO IV. METODOLOGÍA.....	31
4.1 Etapas de desarrollo.....	31
4.1.1 Revisión de la Literatura.....	31
4.1.2 Generación y recopilación de datos.....	32
4.1.2.1 QGSJET-II-04.....	32
4.1.2.2 GHEISHA.....	32
4.1.2.3 CONEX.....	33
4.1.2.4 THIN	33
4.1.2.5 INCLINED	33
4.1.2.6 PLOTSH2	34



4.1.2.6.1	Map2plot.....	35
4.1.2.7	Configuración de la simulación	36
4.1.3	Elección de Redes.....	40
4.1.4	Formación y Evaluación.....	40
4.1.5	Discusión y Análisis de Resultados.....	42
CAPÍTULO V. RESULTADOS Y ANÁLISIS.....		43
5.1	Primer experimento: Conjunto promediado por vista	43
5.2	Segundo experimento: Conjunto promediado general.....	46
5.2.1	Comparación de los conjuntos de datos	50
	Comparación de Desempeño.....	50
CAPÍTULO VI. CONCLUSIONES Y TRABAJO FUTURO		51
6.1	Limitaciones.....	52
6.2	Trabajo Futuro.....	52
	ACRÓNIMOS	56



ÍNDICE DE TABLAS

Tabla 1 Ejemplo de vistas XZ, YZ y XY	35
Tabla 2 Ejemplo Input CORSIKA 7.7550.....	36
Tabla 3 Valores de simulaciones	38
Tabla 4 Métricas (MVTec Software GmbH, 2024).....	40
Tabla 5 Conjunto de datos por vista promediado	44
Tabla 6 Conjunto de datos general promediado	47
Tabla 7 Comparación de desempeño	50

ÍNDICE DE FIGURAS

Figura 1. Comparación gamma-protón (Hampel-Arias, Z., & Westerhoff, S., 2015))	18
Figura 2 Componentes de AERIE	20
Figura 3 Dispersión de cascada gamma/hadrón en un área de 4 m X 4 m. Partícula primaria protón. Pendiente del espectro de partículas primarias -2,0. Rango de energía 5.E9 2. E12 (GeV). Rango del ángulo cenital [0°,65°]. Rango del ángulo azimutal [-180°,1] ...	21
Figura 4 Representación de las EAS ((CTAO), n.d.)	24
Figura 5 Componentes EAS (Glombitza, J. , 2017)	25
Figura 6 Espectro electromagnético (Cherenkov Telescope Array Observatory (CTAO), s.f.).....	27
Figura 7 Flujo de actividades.....	31
Figura 8 Sistema de coordenadas en CORSIKA 7.7550	34

CAPÍTULO I. INTRODUCCIÓN

1.1 Motivación

El estudio de los rayos cósmicos y las cascadas extensas de aire (EAS, por sus siglas en inglés) es un campo de investigación crucial para comprender la física de partículas de alta energía y los procesos astrofísicos que ocurren en el universo. Las EAS son generadas cuando partículas primarias, como protones, muones, núcleos de hierro o rayos gamma, interactúan con la atmósfera terrestre, produciendo una compleja red de partículas secundarias. La correcta clasificación de estas partículas es esencial para inferir información sobre su origen, composición y mecanismos de aceleración, lo que a su vez permite avanzar en la comprensión de fenómenos astrofísicos como los agujeros negros supermasivos, las supernovas y las explosiones de rayos gamma.

Sin embargo, la naturaleza estocástica y la diversidad morfológica de las cascadas presentan desafíos importantes para su clasificación. Las técnicas tradicionales basadas en análisis estadísticos han demostrado ser limitadas en escenarios de alta complejidad, especialmente cuando los datos son generados por simulaciones como las obtenidas con CORSIKA 7.7550, que representan detalladamente los procesos de interacción y decaimiento. Ante este panorama, el uso de redes neuronales convolucionales profundas (CNNs) ofrece un enfoque prometedor para optimizar la clasificación de partículas primarias mediante el análisis de imágenes sintéticas de EAS.

1.2 Justificación

La incorporación de metodologías basadas en inteligencia artificial y aprendizaje profundo en el análisis de cascadas extensas de aire representa un avance significativo en el campo de la astrofísica de partículas. Las redes neuronales convolucionales, en particular, han demostrado ser herramientas poderosas para identificar patrones complejos en imágenes, lo que las hace ideales para abordar los desafíos inherentes al estudio de las EAS.

El desarrollo de una metodología que permita clasificar partículas primarias con precisión no solo mejorará nuestra capacidad para analizar datos simulados, sino que también facilitará la interpretación de datos experimentales en observatorios como HAWC,

Observatorio Pierre Auger o futuros proyectos de detección como SWGO. Además, este enfoque puede contribuir a la optimización de recursos computacionales mediante el diseño de modelos que balanceen precisión y eficiencia, promoviendo su aplicabilidad en diversas configuraciones y plataformas.

Este proyecto también responde a la creciente necesidad de integrar herramientas computacionales avanzadas en la investigación científica, alineándose con tendencias globales que buscan innovar en el análisis de grandes volúmenes de datos y la extracción de conocimiento en áreas interdisciplinarias.

1.3 Objetivos

1.3.1 Objetivo general

Desarrollar y evaluar una metodología basada en redes neuronales convolucionales profundas para la clasificación de cascadas extensas de aire (EAS) generadas por rayos cósmicos, optimizando la identificación de cascadas en imágenes simuladas en CORSIKA 7.7550.

1.3.2 Objetivos específicos

1. Generar y preprocesar un conjunto de imágenes sintéticas de cascadas extensas de aire (EAS) utilizando CORSIKA 7.7550, en vistas XZ, YZ y XY, tanto en color como en escala de grises, asegurando su calidad y adecuación para el entrenamiento de modelos.
2. Entrenar modelos de redes neuronales convolucionales con los datos sintéticos obtenidos, clasificando partículas primarias como hierro, protón, muón positivo y rayos gamma.
3. Evaluar el rendimiento de diferentes arquitecturas de redes neuronales (ResNet-50, Compact y AlexNet) en la tarea de clasificación de partículas primarias, considerando métricas clave como precisión, F1-Score y tiempo de inferencia, identificando fortalezas y debilidades de cada modelo.
4. Interpretar los resultados de clasificación para determinar patrones y limitaciones en la separación de partículas, prestando especial atención a los desafíos asociados con la morfología de las cascadas y los procesos de interacción y decaimiento en la atmósfera.

1.4 Estructura de tesis

La presente tesis es el resultado de un esfuerzo que combina la física de partículas, la astrofísica y la inteligencia artificial, específicamente el aprendizaje profundo. En un contexto donde los rayos cósmicos y las cascadas extensas de aire (EAS) representan una ventana hacia los procesos más energéticos del universo, esta investigación busca contribuir al desarrollo de metodologías para clasificar partículas primarias utilizando simulaciones computacionales y técnicas modernas de análisis de imágenes.

El trabajo aquí presentado se enmarca en el análisis de datos simulados generados por CORSIKA 7.7550, un software líder en simulación de cascadas extensas de aire. Con el apoyo de herramientas de aprendizaje profundo como las redes neuronales convolucionales (CNNs), se explora cómo optimizar la clasificación de partículas como protones, muones positivos, núcleos de hierro y rayos gamma. Este enfoque no solo permite una comprensión más detallada de los fenómenos físicos subyacentes, sino que también establece un marco para futuras aplicaciones en experimentos de detección de rayos cósmicos.

El documento se estructura en capítulos que guían al lector desde los fundamentos teóricos hasta los resultados experimentales y las conclusiones, brindando una visión integral del proceso de investigación. A lo largo del texto, se abordan temas como el estado del arte en la clasificación de partículas, el marco teórico de las cascadas atmosféricas y las metodologías utilizadas, con un enfoque especial en la evaluación del desempeño de diferentes arquitecturas de CNNs.

Espero que este trabajo no solo contribuya al avance del conocimiento en el campo de la astrofísica de partículas, sino que también sirva como punto de partida para investigaciones futuras en la integración de herramientas de inteligencia artificial en la física experimental.

La tesis está organizada en los siguientes capítulos, proporcionando un recorrido lógico desde la motivación del problema hasta los resultados obtenidos y las proyecciones futuras:

- **Capítulo I: Introducción:**
 - Este capítulo define la motivación, justificación y objetivos (general y específicos), además de esbozar la metodología empleada y la estructura general de la tesis.
- **Capítulo II: Estado del Arte:**
 - Presenta una revisión exhaustiva de trabajos previos y desarrollos recientes en simulación de cascadas atmosféricas y técnicas de aprendizaje profundo aplicadas a la física de partículas.
- **Capítulo III: Marco Teórico:**
 - Proporciona los fundamentos teóricos necesarios para entender los conceptos clave, incluyendo el software CORSIKA 7.7550, las interacciones de partículas en la atmósfera, y los principios detrás de las cascadas extensas de aire.
- **Capítulo IV: Metodología:**
 - Detalla el diseño experimental, desde la simulación y preprocesamiento de datos hasta la selección y entrenamiento de redes neuronales convolucionales.
- **Capítulo V: Resultados y Análisis:**
 - Expone los resultados obtenidos en dos experimentos principales: análisis promediado por vista y análisis general. También incluye una comparación detallada del desempeño de los modelos.
- **Capítulo VI: Conclusiones y Trabajo Futuro:**
 - Resume los hallazgos clave, identifica las limitaciones del trabajo y propone posibles líneas de investigación futuras.

Esta estructura tiene como objetivo presentar la investigación de manera clara, lógica y accesible, permitiendo al lector profundizar en cada aspecto del proyecto según su interés o necesidad.

CAPÍTULO II. ESTADO DEL ARTE

2.1 Clasificación con Random Forest

La clasificación precisa de eventos gamma y hadrónicos en telescopios atmosféricos Cherenkov es un desafío fundamental en astrofísica. En este contexto, la investigación realizada por (Sharma, M., Nayak, J., Koul, M. K., Bose, S., & Mitra, A., 2014), representa un avance significativo en el uso de aprendizaje automático para optimizar esta tarea. El estudio se enfoca en mejorar la clasificación de estos eventos utilizando métodos avanzados de clasificación.

En el estudio se evaluaron cinco técnicas principales de aprendizaje automático: Random Forest (RF), Redes Neuronales Artificiales (ANN), Análisis Discriminante Lineal (DISC), Clasificador Naive Bayes (NB) y Máquinas de Soporte Vectorial (SVM). A partir de una base de datos generada mediante simulaciones de Monte Carlo con el código CORSIKA 7.7550, se utilizaron parámetros de imagen de Hillas, como longitud, ancho, distancia y tamaño, para realizar la segregación gamma/protón. El análisis de esta investigación concluyó que **Random Forest (RF)** se destacó como el método más efectivo entre los evaluados, gracias a sus características clave:

- **Rendimiento Superior:** Destaca la sensibilidad en la clasificación de eventos gamma, superando a los otros métodos de aprendizaje automático, incluidos ANN, DISC, NB y SVM. Este rendimiento superior refuerza su capacidad para manejar datos complejos y de alta dimensión.
- **Manejo de Datos Multidimensionales:** El algoritmo RF sobresale en la gestión de conjuntos de datos multidimensionales, como los utilizados en astronomía de rayos gamma, al abordar interdependencias entre parámetros que representan desafíos para métodos lineales convencionales.
- **Robustez y Generalización:** Utiliza múltiples árboles de decisión en combinación con técnicas como el bagging (Bootstrap Aggregating) y la selección aleatoria de características.

- **Comparación con Métodos Convencionales:** En comparación con métodos tradicionales como el supercorte dinámico, RF mostró un poder de clasificación y una sensibilidad significativamente superiores, consolidándose como una herramienta más eficaz para la identificación de eventos gamma.
- **Diversidad de Clasificadores:** Aunque el estudio reconoce que ningún método es universalmente superior, destaca que RF resultó particularmente adecuado para el conjunto de datos específico de esta investigación, gracias a su capacidad para adaptarse a los retos del análisis astrofísico.

2.2 Compactación por Pares

Hampel-Arias, Z., & Westerhoff, S., (2015), en su investigación titulada “Gamma Hadron Separation using Pairwise Compactness Method with HAWC”, implementan el método de Compactación por Pares es significativo en la detección de rayos gamma, en la búsqueda de distinguir entre lluvias de rayos gamma y de rayos cósmicos hadrónicos/núcleos de protón, lo cual es crucial para realizar observaciones precisas en entornos con alto fondo, como los que se encuentran en el Observatorio HAWC. Este método se basa en la técnica de compactación tradicional, analizando la conectividad espacial de una selección de hits (señales de PMT) en el detector.

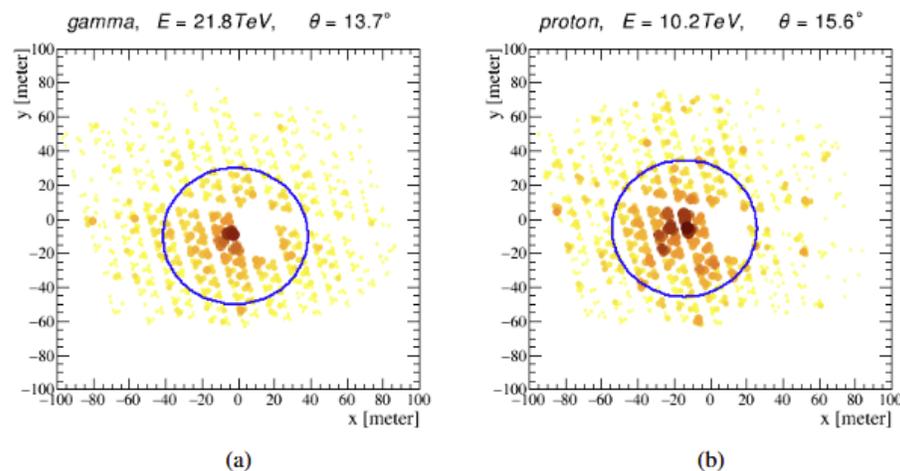


Figura 1. Comparación gamma-protón (Hampel-Arias, Z., & Westerhoff, S., 2015))

Según se menciona el método aprovecha las diferencias en la distribución espacial de los hits entre las cascadas pues se observa que los rayos gamma tienden a tener perfiles laterales más compactos y suaves esto se puede observar en la Figura 1, mientras que las cascadas hadrónicas exhiben fluctuaciones más grandes y más momento transversal, lo que lleva a una mayor dispersión de hits. Al medir la carga media y la separación espacial por pares de los hits seleccionados, el Método de Compactación por Pares puede separar más eficazmente estos dos tipos de cascadas, en caso de las que cuentan un bajo nHit, donde los métodos tradicionales pueden tener dificultades debido a problemas para ajustar con precisión parámetros de la cascada, como la ubicación del núcleo. El Método presentado en esta investigación presenta un buen rendimiento en estos escenarios, aumentando así la sensibilidad general del Observatorio HAWC.

2.2.1 Adaptación de dominio

Drew, R., (2021), incorpora a su investigación el uso de métodos de adaptación de dominio cuyo objetivo es minimizar la discrepancia entre las distribuciones de los dominios fuente y objetivo, lo que permite que el modelo generalice mejor y realice predicciones más precisas aplicadas en el contexto de la separación gamma-hadrones. La tesis explora tanto métodos "superficiales" (como el Pesado de Importancia por Vecinos Más Cercanos, la Discriminación Logística y el Análisis de Componentes de Transferencia), para el proceso de entrenamiento implementa un modelo de red neuronal convolucional (CNN) para la separación gamma-hadrones, mejorando así el rendimiento del modelo más allá de lo que se logró utilizando solo datos simulados.

Destaca que los métodos de adaptación de dominio profunda superan significativamente a los métodos superficiales (como el Pesado de Importancia de Vecinos Más Cercanos y el Análisis de Componentes de Transferencia) en el contexto de la separación gamma-hadrones. Esto indica que los enfoques de aprendizaje profundo pueden capturar patrones y relaciones más complejas en los datos, lo que lleva a una mejor generalización.

2.2.2 Clasificación en HAWC

Para fines prácticos se analiza el software utilizado por el observatorio HAWC (High-Altitude Water Cherenkov) para clasificar cascadas atmosféricas de rayos gamma y rayos cósmicos, está diseñado para analizar los datos obtenidos por sus detectores de Cherenkov en agua. Este software emplea un enfoque basado en algoritmos que procesan señales de partículas cargadas generadas por las cascadas cuando interactúan con los tanques de agua del detector.

Como explica Zúñiga Reyes, A. (2017), el software empleado es AERIE (Analysis and Event Reconstruction Integrated Environment) el cual se utiliza para procesar y analizar cascadas. Codificado como un conjunto de proyectos interconectados en C++, organizados en un framework que gestiona el flujo de datos, almacenamiento y procesamiento.

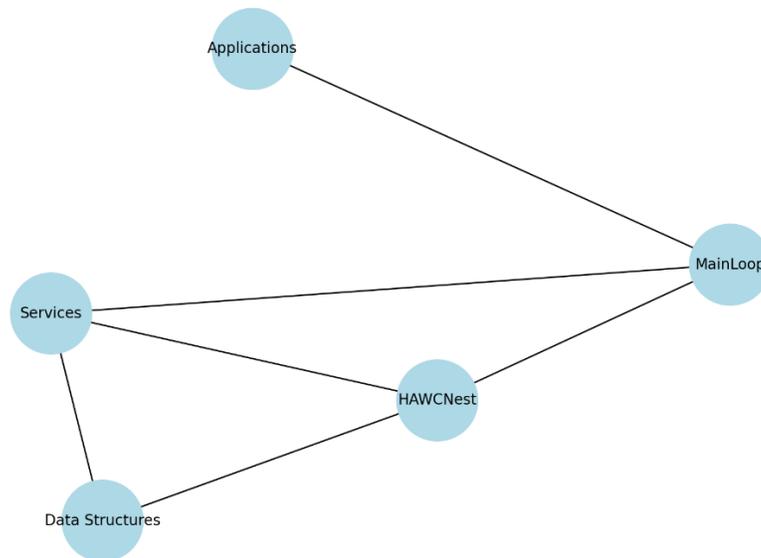


Figura 2 Componentes de AERIE

Como se puede observar en la Figura 2, los componentes principales de AERIE trabajan en conjunto para procesar y analizar datos de manera eficiente. HAWCNest actúa como el núcleo del sistema, registrando e inicializando los servicios necesarios para el procesamiento, aunque no almacena datos directamente. Estos datos son gestionados a través de las Data Structures, que ofrecen una representación en memoria que puede ser modificada dinámicamente por los servicios. Los Services aportan flexibilidad mediante módulos personalizados que incluyen algoritmos para tareas específicas, como cálculos astronómicos

o generación de números aleatorios. El MainLoop organiza el flujo de control del procesamiento, permitiendo la ejecución ordenada de las tareas. Finalmente, las Applications proporcionan herramientas especializadas para análisis básicos, como la creación de mapas celestes y la estimación de espectros energéticos, facilitando la interpretación de los datos procesados. Esta arquitectura modular asegura un enfoque eficiente y adaptable para las necesidades de procesamiento en HAWC. Algo importante a considerar es que durante el proceso de reconstrucción de las cascadas se realizan las simulaciones en CORSIKA 7.7550. Se implementan dos módulos HAWCSim, basado en GEANT4, modela cómo las partículas interactúan con los tanques de agua de HAWC, generando luz Cherenkov que es detectada por los PMTs (fotomultiplicadores) obteniendo lo llamado como “Footprint”, similar al que se muestra a continuación en la Figura 3.

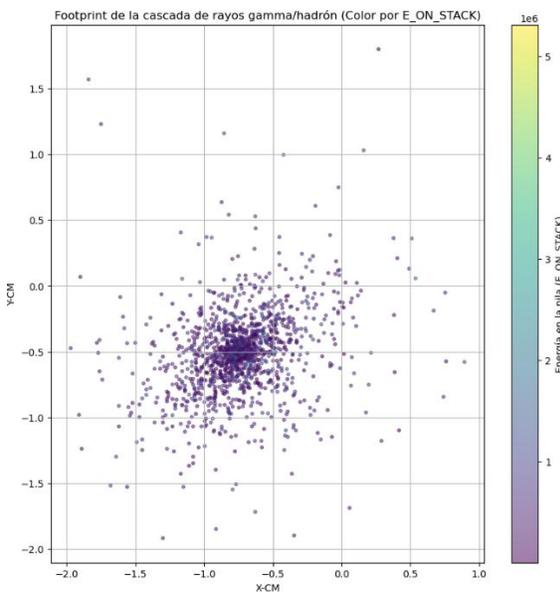


Figura 3 Dispersión de cascada gamma/hadrón en un área de 4 m X 4 m. Partícula primaria protón. Pendiente del espectro de partículas primarias -2,0. Rango de energía $5.E9 \pm 2.E12$ (GeV). Rango del ángulo cenital $[0^\circ, 65^\circ]$. Rango del ángulo azimutal $[-180^\circ, 1^\circ]$

Posteriormente AERIE procesa los datos simulados, agrega ruido y reconstruye eventos para determinar características como energía, dirección y tipo de partícula (rayos gamma o hadrones/protones). Para poder detectar el núcleo se identifica analizando la mayor densidad de energía registrada por los PMTs, así como la dirección de llegada del evento considerando las diferencias en los tiempos de llegada de las señales detectadas por los

PMTs, después de este procedimiento se puede clasificar entre cascadas realizadas por rayos gamma o por protones.

Esta estructura permite a HAWC procesar grandes volúmenes de datos, identificar eventos relevantes y estudiar los rayos gamma con alta precisión, ayudando a explorar fenómenos extremos en el universo.

Arteaga Vázquez, A. (2020), explica que en HAWC, contemplando sus 300 detectores Cherenkov en pleno funcionamiento, registra más de 25,000 eventos por segundo, generando un volumen de datos superior a un terabyte cada día. En su primer año de operaciones, HAWC recopiló aproximadamente 800 mil millones de eventos, lo que equivale a un petabyte de información. De ahí la importancia de acelerar los procesos de clasificación entre los datos de utilidad y los descartables.

CAPÍTULO III. MARCO TEÓRICO

3.1 CORSIKA 7.7550

CORSIKA 7.7550 (*COsmic Ray Simulations for KAscade*) en su versión 7.7500 es un programa de simulación basado en el método matemático y estadístico Monte Carlo para generar cascadas de partículas cósmicas (Extensive Air Showers, EAS) de alta energía, estas partículas al impactar con la atmosfera generan una reacción con los núcleos de aire o, en el caso de las secundarias inestables, se desintegran.

Es una herramienta utilizada en diversos experimentos de rayos cósmicos, abarcando energías desde 10^{12} eV a 10^{20} eV, permite rastrear las trayectorias de partículas secundarias, permitiendo un análisis detallado de las simulaciones.

De acuerdo con el manual (Heck, D., & Pierog, T., 2023), CORSIKA 7.7550 enfrenta retos debido a que los colisionadores actuales no pueden replicar las energías extremas y condiciones de rapidez asociadas con los rayos cósmicos más energéticos. Para superar estas restricciones, el programa incorpora múltiples modelos teóricos de interacción hadrónica, como VENUS, QGSJET y DPMJET, que permiten extrapolar datos a energías y condiciones no observadas experimentalmente. Además, incluye opciones avanzadas para tratar interacciones a bajas energías, logrando así un modelado más completo de las cascadas de partículas.

Las interacciones hadrónicas a energías más bajas se describen mediante las rutinas de interacción GHEISHA, requiriendo la instalación del FLUKA (Ferrari, A., Sala, P. R., Fassò, A., & Ranft, J., 2024) (*FLUktuierende KAskade*) que ayuda a modelar la interacción de partículas con materiales y estructuras, incluyendo fenómenos como dispersión, absorción, producción de partículas secundarias, y transporte de radiación electromagnética y hadrónica, igualmente incorpora el modelo microscópico UrQMD (*Ultra-relativistic Quantum Molecular Dynamics*) (Bleicher, M., Zabrodin, E., Spieles, C., Bass, S. A., Ernst, C., Soff, S., ... & Stöcker, H., 1999) es un modelo computacional ampliamente utilizado para simular interacciones entre partículas a energías altas y ultra-relativistas. Es especialmente valioso en estudios de colisiones hadrónicas, interacciones nucleares y formación de plasmas de quarks y gluones.

3.2 Interacción de las partículas

Cuando una partícula de alta energía, como un núcleo de rayo cósmico o un fotón, penetra las capas superiores de la atmósfera terrestre, colisiona con átomos y moléculas del aire, dando lugar a una cascada de partículas secundarias conocida como EAS. La composición de esta cascada depende del tipo de partícula inicial: si la partícula primaria es un fotón, electrón o positrón, se forma una cascada principalmente electromagnética. En cambio, si la interacción involucra un núcleo de rayo cósmico mediante la fuerza fuerte, además de componentes electromagnéticos, se generan subcascadas de naturaleza hadrónica como se muestra en la Figura 4.

Aunque se les llama “rayos”, los rayos cósmicos son partículas atómicas estándar. Están compuestos en un 99% por protones y núcleos de helio, y el 1% restante por núcleos más pesados, electrones, muones y otras partículas. A pesar de ser partículas “normales”, los rayos cósmicos son especiales porque se aceleran a energías extraordinariamente altas en el cosmos, viajando muy cerca de la velocidad de la luz. Los rayos cósmicos bombardean constantemente la Tierra, pero a pesar de una búsqueda de un siglo, sabemos muy poco sobre sus fuentes de origen y el papel que desempeñan en nuestra propia galaxia y más allá. Esto se debe a que los rayos cósmicos están cargados eléctricamente y, por lo tanto, sus caminos están desordenados en los campos magnéticos entre sus fuentes y la Tierra, lo que hace imposible rastrearlos hasta su origen.

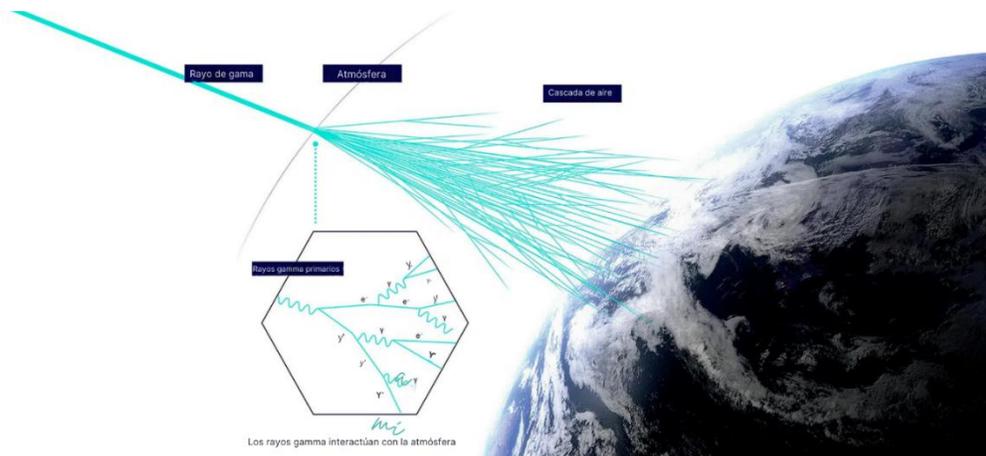


Figura 4 Representación de las EAS ((CTAO), n.d.)

3.3 Cascadas Atmosféricas de Aire

Cuando una partícula primaria de alta energía ingresa a la atmósfera terrestre, interactúa con las moléculas de aire mediante la fuerza nuclear fuerte, dando lugar a la generación de partículas secundarias, predominantemente mesones (como piones) y, en menor medida, bariones. Estas partículas secundarias pueden interactuar nuevamente con las moléculas de aire, desencadenando una cascada de colisiones que produce nuevas generaciones de mesones (Figura 5). (Glombitza, J. , 2017).

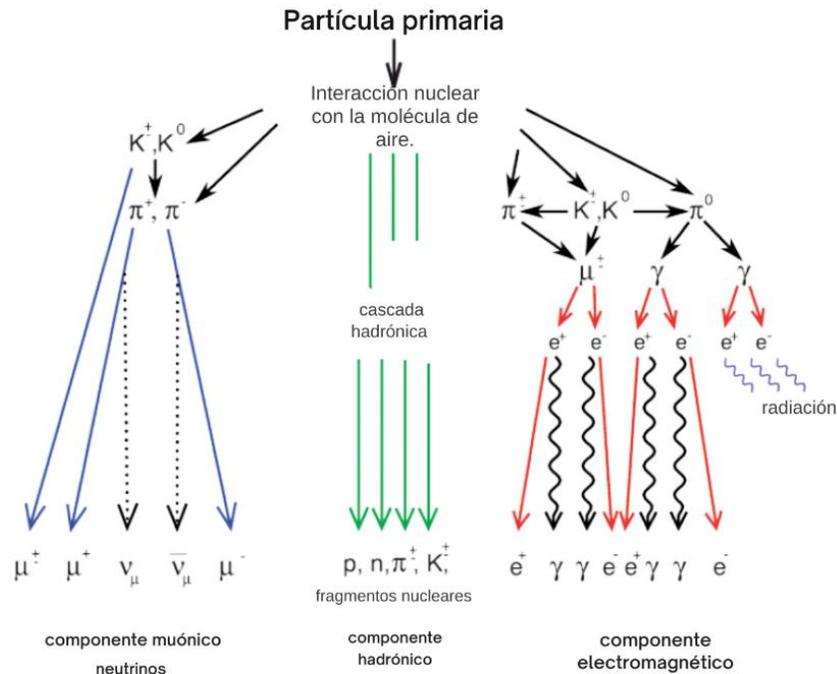


Figura 5 Componentes EAS (Glombitza, J. , 2017)

A medida que estas interacciones progresan, las partículas generadas se dividen en diferentes tipos, cada una transportando una fracción específica de la energía inicial. Este proceso da lugar a la formación de tres componentes principales de las cascadas extensas de aire (Extensive Air Showers, EAS): el componente hadrónico, compuesto principalmente por nucleones y partículas cargadas; el componente muónico, derivado del decaimiento de los mesones; y el componente electromagnético, constituido por fotones, electrones y positrones (Engel, R., Heck, D., & Pierog, T., 2011).

Este fenómeno es esencial en el estudio de los rayos cósmicos, ya que permite interpretar las propiedades de las partículas primarias y las interacciones de alta energía que

ocurren en condiciones extremas. Herramientas como CORSIKA 7.7550 y otros programas de simulación basados en Monte Carlo son fundamentales para modelar estos procesos y analizar los tres componentes principales de las cascadas de partículas, proporcionando una mejor comprensión de la naturaleza y origen de los rayos cósmicos.

3.4 Rayos gamma (rayos- γ) como partículas primarias

Los rayos- γ no están cargados eléctricamente y, por lo tanto, su camino directo hacia la Tierra permite que los rayos gamma transporten información sobre sus fuentes y las partículas relativistas que los crearon.

Como destaca (Cherenkov Telescope Array Observatory (CTAO), s.f.), estas partículas relativistas, se producen cuando los rayos cósmicos interactúan con la materia, los campos magnéticos o la luz en las fuentes o alrededor de ellas, o en el espacio interestelar. Los rayos gamma viajan a través del Universo hasta las galaxias que se encuentran más allá, transportando consigo los secretos de su lugar de origen.

Los fotones de rayos- γ poseen propiedades similares a los fotones de luz visible y las ondas de radio: se desplazan en línea recta, viajan a la velocidad de la luz y no se ven afectados por los campos magnéticos intergalácticos en el espacio exterior. Gracias a estas características, los rayos gamma actúan como sondas precisas, apuntando directamente hacia sus fuentes emisoras. Esto permite explorar regiones remotas del cosmos e incluso áreas oscuras dentro de nuestra propia galaxia, la Vía Láctea. (Zuñiga Reyes, A., 2017)

Es importante destacar que los rayos X y los rayos gamma provenientes de fuentes distantes pueden experimentar dispersiones temporales debido a la interacción con objetos astronómicos intermedios, como los que generan lentes gravitacionales. Estos efectos pueden modificar la trayectoria o el tiempo de llegada de los fotones, aportando información valiosa sobre las estructuras que se encuentran entre la fuente y el observador.

Según explica Zúñiga Reyes, A. (2017), Para caracterizar la radiación gamma, no se utilizan parámetros como la longitud de onda (λ) o la frecuencia (ν), sino su energía (E). Esto se debe a que los fotones de rayos gamma tienen longitudes de onda extremadamente pequeñas, del orden de 10^{-11} metros o incluso menores, lo que los sitúa en el rango de energías muy altas, según la relación de Planck:

$$E = h\nu = \frac{hc}{\lambda}$$

Donde:

- h es la constante de Planck.
- c es la velocidad de la luz.

Estas energías corresponden a cientos de kilo-electronvoltios (keV) o más. La radiación electromagnética con energías superiores a 0.5 MeV se clasifica como rayos gamma, el límite superior corresponde a 10^{20} eV como se puede observar en la Figura 6, siendo estas las partículas más energéticas detectadas en los rayos cósmicos. En la actualidad, se han llegado a registrar rayos gamma con energías que alcanzan decenas de tera-electronvoltios (TeV).

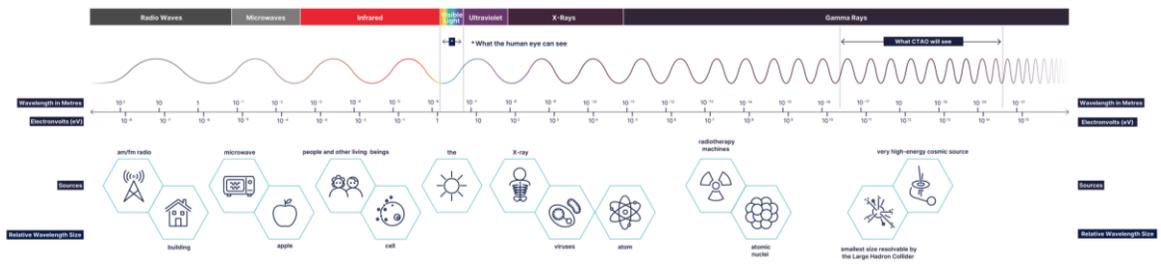


Figura 6 Espectro electromagnético (Cherenkov Telescope Array Observatory (CTAO), s.f.)

Este enfoque en la energía permite describir con mayor precisión el comportamiento y las propiedades de los rayos gamma, que representan el extremo más energético del espectro electromagnético y resultan fundamentales para explorar fenómenos extremos en el universo.

3.4.1 MVTec Deep Learning Tool 24.05.1

La herramienta de Deep Learning Tool de MVTec es un software diseñado para la preparación de datos de imagen y para facilitar flujos de trabajo impulsados por datos en el despliegue de aplicaciones de visión artificial que utilizan técnicas de aprendizaje profundo. Esta herramienta permite gestionar conjuntos de datos de manera sencilla y ofrece etiquetado asistido por inteligencia artificial, lo que facilita el uso de imágenes para el entrenamiento de

redes neuronales. Además, posibilita la evaluación directa de las redes entrenadas. Permite abordar diversas tareas de visión por computadora, tales como clasificación, reconocimiento de objetos, OCR profundo, detección de anomalías en contextos globales y diferentes tipos de segmentación (MVTec Software GmbH, 2024).

- **Etiquetado de Datos**

El etiquetado de datos es uno de los pasos principales en los proyectos de aprendizaje profundo, ya que permite incorporar información sobre cómo resolver un problema específico (MVTec Software GmbH, 2024). Algunos de los métodos incorporados en Deep Learning Tool son:

- Clasificación: En este método, se importan imágenes y se les asigna una clase.
- Detección de Objetos: Se lleva a cabo etiquetando al dibujar rectángulos alrededor de los objetos de interés y asignándoles las clases correspondientes.
- Segmentación: Este proceso consiste en dibujar formas poligonales o pintar máscaras de píxeles alrededor de los objetos relevantes.
- Entrenamiento de OCR Profundo: Se puede aumentar la precisión del reconocimiento óptico de caracteres (OCR) al reentrenar modelos con conjuntos de datos etiquetados de forma eficiente, utilizando sugerencias automáticas para facilitar el proceso.
- Detección de Anomalías: Este tipo de etiquetado implica importar imágenes y clasificarlas como “buenas” o “anomalías”.

- **Entrenamiento**

Durante el proceso de entrenamiento, un modelo previamente entrenado (o preentrenado) es ajustado utilizando un conjunto de datos de imágenes etiquetadas. En cada iteración sobre el conjunto de datos de entrenamiento, el modelo evalúa sus predicciones en comparación con los resultados esperados del conjunto de validación. Este proceso permite ajustar los pesos de la red neuronal, optimizando progresivamente su rendimiento para las siguientes iteraciones. El objetivo es que el modelo generalice mejor los patrones presentes en los datos y mejore su capacidad de realizar predicciones precisas.

De acuerdo con MVTec Software GmbH (2024), en las herramientas modernas de aprendizaje profundo, los usuarios pueden personalizar los parámetros críticos directamente

desde la interfaz de configuración de entrenamiento. Esto incluye la selección de divisiones de datos (entrenamiento, validación y prueba), el ajuste de hiperparámetros como la tasa de aprendizaje, y la elección de las métricas de evaluación. Una vez configurado, el proceso de entrenamiento se puede iniciar, y los usuarios tienen acceso a gráficos y estadísticas que muestran el progreso del modelo y su desempeño en tiempo real.

Algunos de los modelos preentrenados con los que el software permite trabajar son:

- Compact: Compact es una red sencilla y eficiente, diseñada para maximizar su desempeño en tareas específicas, especialmente en dispositivos con hardware limitado. Presenta ventajas para aplicaciones en tiempo real que necesitan rapidez y bajo consumo de recursos. Sin embargo, su capacidad para resolver problemas más complejos es limitada, y no alcanza el nivel de precisión de modelos más avanzados cuando se trabaja con grandes volúmenes de datos.
- Enhanced: Enhanced es una versión mejorada de redes convencionales, diseñada para ser más precisa y ajustarse a problemas específicos. Aunque es muy efectiva en su área, necesita ser optimizada cuidadosamente para cada tarea que se le asigne, lo que puede hacerla menos práctica en situaciones más generales.
- ResNet 50: ResNet 50 es una red con 50 capas que usa conexiones residuales para resolver problemas como el desvanecimiento del gradiente, efectiva en tareas complejas de visión por computadora.
- ResNet 18: ResNet 18 es una versión más ligera y rápida de ResNet, con solo 18 capas. Empleada para tareas menos complejas, donde un balance entre eficiencia y rendimiento es clave. Sin embargo, no alcanza el nivel de detalle y precisión que ofrece ResNet 50 cuando se trata de problemas más complejos.
- MobileNet V2: MobileNet V2 es una red diseñada para dispositivos móviles, pensada para ser eficiente y funcionar bien con recursos limitados. Su diseño innovador la hace rápida, pero no tiene la misma precisión que modelos más avanzados como ResNet 50 y no es adecuada para tareas altamente especializadas o complejas.
- AlexNet: AlexNet es una red usada desde 2012 y famosa por revolucionar la clasificación de imágenes en el desafío ImageNet. Aunque es menos precisa y eficiente que las redes modernas, sigue siendo una opción sólida para tareas de clasificación con conjuntos de datos pequeños o medianos.

- **Evaluación**

La etapa de evaluación es un paso fundamental en el desarrollo de modelos de visión por computadora, ya que permite medir su desempeño utilizando el conjunto de datos de prueba. Esta fase proporciona información clave sobre la capacidad del modelo para generalizar y resolver problemas en escenarios prácticos. Para los especialistas en visión artificial, esta evaluación es esencial para identificar fortalezas, debilidades y oportunidades de mejora en el modelo.

La herramienta de evaluación ofrece un entorno completo para analizar el rendimiento del modelo entrenado (MVTec Software GmbH, 2024). Entre las funcionalidades destacadas se encuentra la generación de mapas de calor, que permiten visualizar de manera intuitiva las clases predichas para cada imagen procesada. Estos mapas son especialmente útiles para identificar patrones en las predicciones y áreas problemáticas.

Otra característica relevante es la matriz de confusión interactiva, que ayuda a detectar y analizar errores de clasificación. Esta herramienta no solo muestra dónde se encuentran los errores, sino que también facilita la comprensión de las relaciones entre las clases y las áreas donde el modelo tiene mayor dificultad.

CAPÍTULO IV. METODOLOGÍA

4.1 Etapas de desarrollo

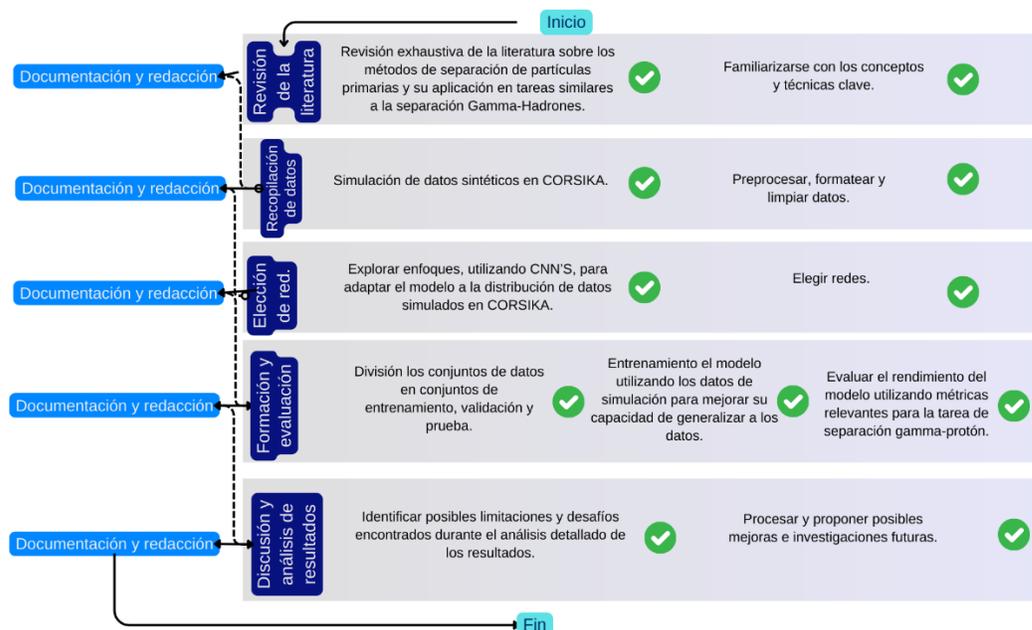


Figura 7 Flujo de actividades

A continuación, se relata el flujo de actividades y etapas para el desarrollo de los experimentos y la metodología empleada:

4.1.1 Revisión de la Literatura

Se realizó un análisis detallado y exhaustivo de los métodos existentes para la separación de partículas primarias generadas en cascadas extensas de aire (EAS), aplicados en diferentes observatorios de altas energías con un enfoque particular en las aplicaciones del aprendizaje profundo para la clasificación de las cascadas producidas partículas primarias, para ello se han seleccionado la partícula de protón, muón positivo, hierro y gamma para las cuales se estudió su interacción en la atmosfera terrestre, procesos de decaimiento y eliminación, así mismo se estudió el sitio de observatorio HAWC para tomar como referencia en las futuras simulaciones.

Se examinó el estado del arte en el uso de CORSIKA 7.7550 para la simulación de datos sintéticos, identificando enfoques relevantes y limitaciones en trabajos previos,

creando una base teórica sólida que permita entender las técnicas actuales y su aplicabilidad al problema planteado, así como definir áreas de mejora e innovación para las metodologías propuestas.

4.1.2 Generación y recopilación de datos

Para esto se empleó el programa CORSIKA 7.7550 para simular interacciones y desintegraciones de núcleos, hadrones, muones, electrones y fotones, dentro la simulación permite configurar las variables iniciales para obtener la energía, la ubicación, la dirección y los tiempos de llegada de todas las partículas secundarias.

Las rutinas principales de CORSIKA 7.7550 se encuentran codificadas en FORTRAN estándar, sin embargo, posteriormente los archivos pueden ser convertidos para trabajarlos en C++ o Python, es necesario ejecutar las rutinas en un sistema operativo con base UNIX, para efectos de esta tesis se implementó Ubuntu 22.04.5.

CORSIKA 7.7550 trabaja con un conjunto de rutinas cuyo diseño es modular de esta manera organiza los datos de entrada y salida, rastrea las trayectorias de partículas y considera factores importantes como la pérdida de energía por ionización y los efectos del campo magnético terrestre (Engel, R., Heck, D., & Pierog, T., 2011).

A continuación, se mencionan los modelos y herramientas implementados para las simulaciones de las cascadas.

4.1.2.1 QGSJET-II-04

QGSJET-II-04 es un modelo utilizado en simulaciones para describir colisiones entre hadrones y núcleos, así como entre núcleos a altas energías, QGSII selecciona las rutinas específicas de QGSJET-II-04,

4.1.2.2 GHEISHA

GHEISHA 2002d es un modelo de bajas energías que simula las colisiones hadrón-núcleo, el cual se combina con QGSJET-II-04, ofreciendo flexibilidad en la simulación de cascadas atmosféricas.

4.1.2.3 CONEX

CONEX es una herramienta que se encarga de calcular, de forma más eficiente el desarrollo de las cascadas atmosféricas después de las primeras interacciones de partículas. Es necesario combinarlo con modelos de interacción de alta energía como EPOS, QGSJET, QGSII o SIBYLL, junto con cualquiera de los modelos de baja energía disponibles. Además, requiere activar ciertas opciones del programa, como CURVED, SLANT, THIN y UPWARD, para garantizar que las simulaciones sean precisas y optimizadas. (Heck, D., & Pierog, T., 2023).

4.1.2.4 THIN

THIN activa un mecanismo de reducción (thinning) que ayuda a disminuir el tiempo de cálculo en simulaciones completas cambiándose con CONEX, para las simulaciones incorpora el efecto LPM (Landau-Pomeranchuk-Migdal), que afecta las interacciones electromagnéticas a muy altas energías.

Como efecto secundario al implementar esta función se reduce el número de partículas simuladas en cascadas atmosféricas al seleccionar un subconjunto representativo de las mismas, asignando a cada una un peso que refleja el número de partículas que representan. Esto permite que se modele las cascadas con un menor costo computacional, manteniendo resultados estadísticamente significativos.

4.1.2.5 INCLINED

INCLINED es una herramienta que forma parte del paquete COAST (CORSIKA 7.7550 Advanced Simulation Tool), que registra las partículas que llegan al suelo en un plano inclinado, en lugar de un plano horizontal tradicional, manejar geometrías más complejas, útil en situaciones donde los detectores están ubicados en regiones montañosas como lo es el caso del observatorio HAWC, la herramienta permite afectar la distribución de partículas al llegar al suelo.

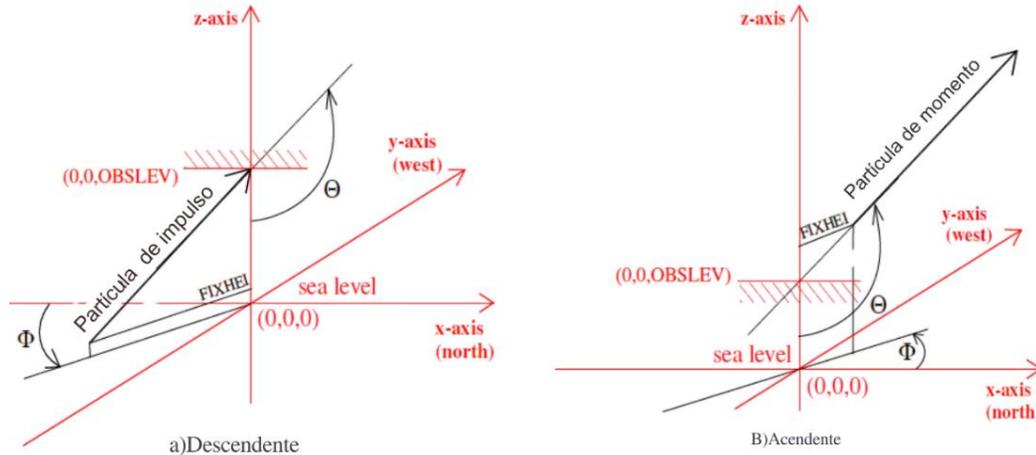


Figura 8 Sistema de coordenadas en CORSIKA 7.7550

Se utiliza un sistema de coordenadas como el que se muestra en la Figura 8 para describir las trayectorias de las partículas y simular su movimiento. El ángulo cenital (θ) mide la inclinación de la trayectoria de la partícula respecto al eje vertical ($-Z$). Si $\theta = 0^\circ$, la partícula se mueve directamente hacia abajo. Si $\theta > 90^\circ$, la partícula está subiendo, es decir, se trata de una cascada ascendente. En el caso del ángulo azimutal (ϕ) mide la dirección horizontal de la partícula, comparando su trayectoria con el eje X (que apunta al Norte, midiendo en sentido antihorario). (Heck, D., & Pierog, T., 2023)

Dependiendo de la parte las herramientas usadas, las definiciones de los ejes pueden variar:

- En EGS4, el eje Z apunta hacia abajo, el eje Y hacia el Este, y X sigue apuntando al Norte.
- En CONEX, el eje X apunta al Este, Y apunta al Norte, y Z permanece igual.

4.1.2.6 PLOTSH2

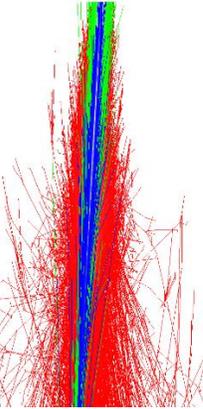
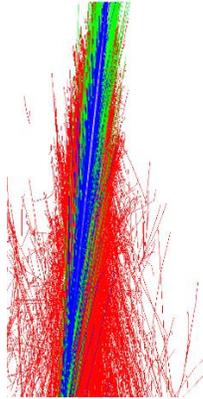
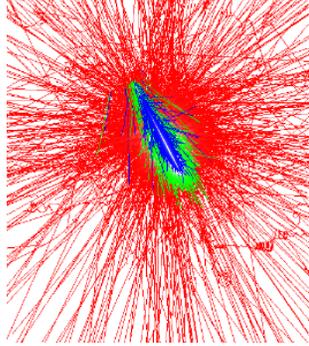
PLOTSH2 es una herramienta que permite generar archivos .map y .track con las trayectorias de las partículas electromagnéticas (electrones, positrones y fotones), muónicas y hadrónicas (protones, neutrones y otros núcleos) que participan en el desarrollo de una cascada atmosférica.

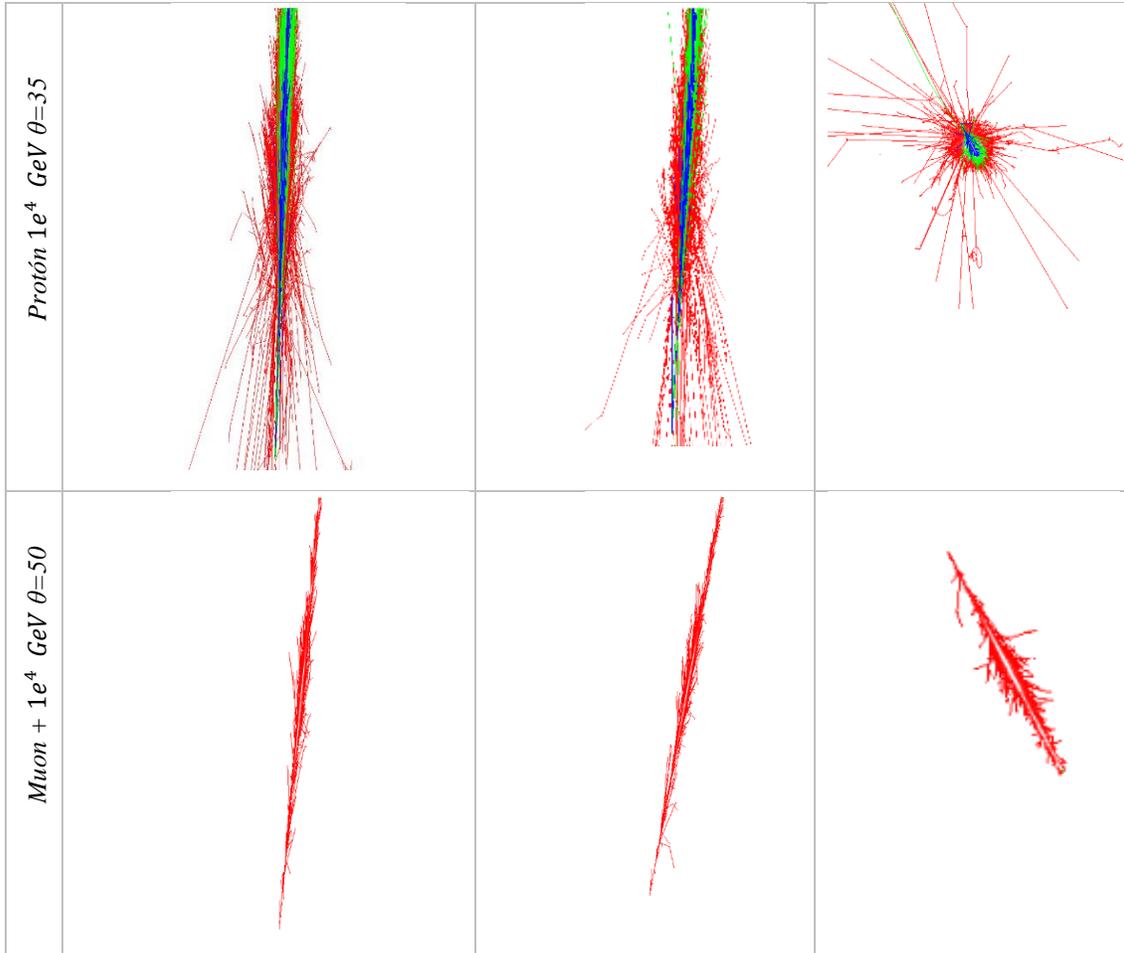
4.1.2.6.1 Map2plot

Crear representaciones gráficas a partir de los datos de simulación, específicamente para los datos de cascadas de partículas generadas por el código CORSIKA 7.7550. Ayuda a convertir la salida de simulaciones en representaciones visuales que pueden ser analizadas e interpretadas más fácilmente arrojando como resultado los archivos .map y .track.

Con esto se permite generar imágenes en vistas XZ, YZ y XY, tanto en color como en escala de grises, asegurando que cada conjunto capture características clave de las cascadas, algunas de las imágenes adquiridas se muestran a continuación en la Tabla 1.

Tabla 1 Ejemplo de vistas XZ, YZ y XY

	XZ	YZ	XY
Gamma $1e^2$ $GeV\theta=0$			
Hierro $1e^6$ $GeV\theta=35$			



4.1.2.7 Configuración de la simulación

A continuación, en la Tabla 2 se muestra el ejemplo de input para CORSIKA 7.7550 usado con los modelos anteriormente mencionados:

Tabla 2 Ejemplo Input CORSIKA 7.7550

Descripción	Variable	Valor
Número de simulación	RUNNR	1
Número de cascadas generadas	NSHOW	1
Partícula primaria	PRMPAR	1 = gamma, 14 = protón, ...
Pendiente espectral	ESLOPE	-2.7
Rango energético de la partícula primaria (GeV)	ERANGE	[m, n]

Rango de ángulo cenital (grados)	THETAP	35. 35
Rango de ángulo azimutal (grados)	PHIP	-180. 180
Secuencia de números aleatorios	SEED	(definido por el usuario)
Nivel de observación (en cm)	OBSLEV	4.10E+04
Altitud inicial (g/cm ²)	FIXHEI	0. 0
Campo magnético	MAGNET	27.2 14.0
Parámetros de reducción	THIN	1e-6 10. 0.
Cortes energéticos para partículas	ECUTS	1. 1. 1.e-3 1.e-3
Gráficos de desarrollo de cascada	PLOTSH	T
Información adicional para muones	MUADDI	T
Ángulo de dispersión múltiple de muones	MUMULT	T
Bandera de interacción electromagnética (NKG, EGS)	ELMFLG	F T
Factor de longitud de paso para dispersión múltiple	STEPFC	1
Radio externo para distribución latitudinal NKG	RADNKG	2.00E+04
Distribución longitudinal, paso, ajuste y salida	LONGI	T 10. F T
Corte en el factor gamma para impresión	ECTMAP	1.00E+05
Número máximo de eventos impresos	MAXPRT	1
Directorio de salida	DIRECT	/
Generar archivo .dbase	DATBAS	F
Usuario	USER	veronica
Bandera de depuración y unidad de registro	PAROUT	T
Depuración y log (unidad, frecuencia, etc.)	DEBUG	F 6 F 1000000
Termina la entrada	EXIT	

Los cambios entre las partículas simuladas difieren entre el valor del rango de energía y la variación del ángulo cenital, los valores se muestran en la Tabla 3.

Tabla 3 Valores de simulaciones

		ERANGE	THETAP
HIERRO (5626 EN CORSIKA 7.7550)	$1e^4$		$\Theta=0$
			$\Theta=20$
			$\Theta=35$
			$\Theta=50$
	$1e^5$		$\Theta=0$
			$\Theta=20$
			$\Theta=35$
			$\Theta=50$
	$1e^6$		$\Theta=0$
			$\Theta=20$
			$\Theta=35$
			$\Theta=50$
PROTÓN (14 EN CORSIKA 7.7550)	$1e^7$		$\Theta=0$
			$\Theta=20$
			$\Theta=35$
			$\Theta=50$
	$1e^2$		$\Theta=0$
			$\Theta=20$
			$\Theta=35$
			$\Theta=50$
$1e^3$		$\Theta=0$	
		$\Theta=20$	
		$\Theta=35$	
		$\Theta=50$	
$1e^4$		$\Theta=0$	
		$\Theta=20$	
		$\Theta=35$	
		$\Theta=50$	

MUÓN + (6 EN CORSIKA 7.7550)	$1e^5$	$\Theta=0$
		$\Theta=20$
		$\Theta=35$
	$1e^2$	$\Theta=0$
		$\Theta=20$
		$\Theta=35$
		$\Theta=50$
	$1e^3$	$\Theta=0$
		$\Theta=20$
		$\Theta=35$
		$\Theta=50$
	GAMMA (1 EN CORSIKA 7.7550)	$1e^4$
		$\Theta=20$
		$\Theta=35$
		$\Theta=50$
$1e^5$		$\Theta=0$
		$\Theta=20$
		$\Theta=35$
$1e^2$		$\Theta=0$
		$\Theta=20$
		$\Theta=35$
		$\Theta=50$
		$\Theta=0$
	$\Theta=20$	
	$\Theta=35$	
	$\Theta=50$	
	$\Theta=0$	
	$\Theta=20$	
	$\Theta=35$	
	$\Theta=50$	
	$\Theta=0$	
	$\Theta=20$	
	$\Theta=35$	
	$\Theta=50$	

Posteriormente al evaluar la calidad de las imágenes generadas mediante métricas específicas para garantizar su utilidad en el entrenamiento, se procede a generar un conjunto de datos estructurado que represente fielmente las características de las cascadas extensas de aire, preparado para ser utilizado en el entrenamiento de las redes neuronales.

4.1.3 Elección de Redes

Al identificar y configurar arquitecturas de redes neuronales convolucionales incorporadas a Deep Learning Tool 24.05.1, se evaluó por su historia en proyectos similares implementado en la clasificación de las imágenes tomando en cuenta el rendimiento como la eficiencia computacional.

Considerando las características de la computadora a disposición: sistema operativo Windows 11 con NVIDIA GeForce RTX 4060 Laptop GPU (gpu:0), procesador Intel(R) Core (TM) i9-14900HX a 2.20 GHz.

Dentro del software Deep Learning Tool 24.05.1, se analizan las arquitecturas de Compact, AlexNet y ResNet50, estudiando su capacidad para procesar datos con alta complejidad morfológica como las de las cascadas, se adaptan las configuraciones de las redes para ajustarse a las características del conjunto de datos, como el tamaño de las imágenes y las vistas generadas y posteriormente se realiza una comparación teórica inicial entre los modelos para anticipar ventajas y limitaciones, como el tiempo de inferencia, la capacidad de generalización y la sensibilidad a los cambios en las condiciones de entrada.

4.1.4 Formación y Evaluación

Al entrenar las redes neuronales seleccionadas utilizando los datos simulados en CORSIKA 7.7550, evaluando su rendimiento con métricas específicas que se muestran en la Tabla 4, las cuales permitan comparar su efectividad en la tarea de clasificación de partículas primarias.

Tabla 4 Métricas (MVTec Software GmbH, 2024)

Métrica	Significado
---------	-------------

Top1-error	Medida de evaluación que representa la media armónica de la precisión y la recuperación. Proporciona una forma de resumir tanto la precisión (la proporción de verdaderos positivos sobre el total de positivos) como la recuperación (la proporción de verdaderos positivos sobre el total de positivos reales) en un solo número. Útil en situaciones donde hay un desequilibrio en las clases, ya que equilibra ambas métricas y resalta el rendimiento del modelo en términos de eficacia al clasificar eventos positivos.
Perdida	Función que se optimiza durante el proceso de entrenamiento, mide si las predicciones del modelo se alinean con los valores reales en los datos de entrenamiento, compara las predicciones del modelo con la información objetiva que se espera encontrar y calcula una penalización por las desviaciones entre ambas.
Inferencia por imagen	Tiempo de proceso de obtener predicciones a partir de un modelo entrenado usando imágenes nuevas, evalúa el rendimiento del modelo en datos no utilizados en el entrenamiento.
Tiempo de procesamiento	Tiempo que toma al sistema para procesar una sola imagen a través del modelo.
Tiempo total	Suma del tiempo de procesamiento, inferencia, otros tiempos asociados, como la carga de datos y la preparación del modelo.
Precisión	Es la proporción de predicciones verdaderamente positivas sobre el total de predicciones positivas.
F1-score	Mide el equilibrio armónico entre la precisión y el recall de un modelo de clasificación en un solo número. Un F1-score alto indica que un modelo tiene un buen equilibrio entre precisión y recall.

<p>Recall</p>	<p>Mide la proporción de verdaderos positivos (TP) respecto a todos los positivos reales (TP y falsos negativos, FN). Se calcula con la siguiente fórmula:</p> $Recall = \frac{TP}{TP + FN}$ <p>Un alto valor de recall sugiere que el modelo es efectivo para detectar los positivos, aunque puede que también incremente el número de falsos positivos.</p>
<p>Mejor época</p>	<p>Se refiere a la iteración del entrenamiento en la que el modelo logra el mejor rendimiento.</p>

Se procede a dividir el conjunto de datos en entrenamiento (70%), validación (15%) y prueba (15%), asegurando una distribución adecuada y balanceada.

Se entrena cada modelo utilizando el conjunto de entrenamiento, ajustando hiperparámetros como el tamaño del lote (batch size), tasa de aprendizaje (learning rate) y número de épocas para maximizar la precisión y minimizar la pérdida.

Se realiza una 5 iteración de entrenamiento y validación para promediar los valores obtenidos, esto como un conjunto de pruebas y medir el desempeño real de los modelos en datos no vistos, las tablas con los valores promediados se encuentran en la Tabla 5 y Tabla 6.

4.1.5 Discusión y Análisis de Resultados

Se procede a interpretar los resultados obtenidos en las evaluaciones para identificar patrones, limitaciones y áreas de mejora en la metodología propuesta.

Se analizar los patrones de clasificación logrados por cada modelo desglosándolo en una tabla, esto permite examinar los factores que afectan el rendimiento, como la influencia de las vistas (XZ, YZ, XY) y la escala de las imágenes (color vs. grises).

Se proponen mejoras o entrenamiento con datos más diversos, así como se logra plantear líneas futuras de investigación, incluyendo la integración de los modelos en experimentos reales o la optimización de su desempeño en plataformas en observatorios específicos.

CAPÍTULO V. RESULTADOS Y ANÁLISIS

5.1 Primer experimento: Conjunto promediado por vista

El entrenamiento del modelo se realizó considerando el 70% de las imágenes totales (42 imágenes) para entrenar los modelos, permitiendo que aprendieran las características relevantes de las partículas. Un 15% de las imágenes (9 imágenes) se empleó para ajustar los hiperparámetros y evitar el sobreajuste del modelo. El 15% restante de las imágenes (9 imágenes) se usó para evaluar el desempeño final del modelo en datos no vistos durante el entrenamiento.

La ejecución y el entrenamiento se llevaron a cabo en una NVIDIA GeForce RTX 4060 Laptop GPU. Esto permitió acelerar el procesamiento de los datos y las operaciones matemáticas intensivas, especialmente las relacionadas con las redes neuronales convolucionales.

Se utilizó el optimizador Adam, conocido por combinar las ventajas de los métodos de gradiente descendente estocástico adaptativo logrando una convergencia más rápida y estable.

Hiperparámetros principales:

- **Batch Size:** Se procesaron lotes de 32 imágenes en cada iteración, lo cual es un valor común para balancear el uso de memoria y la estabilidad del gradiente.
- **Iteraciones:** Se realizaron 100 iteraciones totales para cada época.
- **Épocas:** El modelo fue entrenado durante 50 épocas, proporcionando suficientes oportunidades para ajustar los pesos y minimizar la función de pérdida.
- **Learning Rate:** Se utilizó una tasa de aprendizaje de 0.0001, un valor bajo que favorece ajustes precisos de los pesos del modelo.

Tabla 5 Conjunto de datos por vista promediado

	XY COLOR		YZ COLOR		XZ COLOR			
	Alex net	Compact	Resnet 50	Alex net	Compact	Resnet 50	Alex net	Compact
Top1-Error	5.222	6.778	3.3325	3.7475	5	0	0	0.334
Pérdida	3.37E-06	4.69E-04	4.83E-05	1.22E-04	3.14E-03	3.51E-05	1.91E-04	6.02E-04
Inferencia por imagen ms	1.022	0.982	31.466	0.69	0.484	51.746	0.458	0.648
Tiempo de procesamiento * ms	0.776	0.676	1.112	0.91	0.824	1.112	0.936	0.798
Tiempo total* ms	1.8	1.652	32.578	1.604	1.304	52.858	1.396	1.444
Exactitud %	95.14	93.222	97.08	96.336	95	99.44333333	100	98.33
F1-Score %	94.278	92.498	97.35	96.2575	94.992	100	100	99.165
Precisión %	95.716	93.56	97.612	96.458	95.114	99.48	100	98.96
Recall %	94.5	92.666	97.334	96.336	95	100	100	99.165
Mejor época	24	37.8	42.4	42.6	38.4	33	42.8	39.4

	XY GRIS				YZ GRIS				XZ GRIS				
	Resnet 50	Alex net	Compact	Resnet 50	Alex net	Compact	Resnet 50	Alex net	Compact	Resnet 50	Alex net	Compact	Resnet 50
	6.668	3.664	4.816666667	4.165	4.332	10.332	3.886666667	3.3325	11.334	18.3325			
	6.72E-05	4.80E-05	5.23E-04	4.67E-05	1.50E-04	1.24E-03	1.19E-04	1.43E-04	7.89E-04	4.24E-05			
	49.33	0.47	0.504	48.522	0.412	0.47	30.376	0.638	0.426	106.284			
	0.538	0.5	0.442	0.546	0.442	0.436	0.644	0.544	0.422	1.53			
	49.866	0.97	0.95	49.064	0.862	0.908	31.018	1.18	0.844	107.818			
	93.332	96.336	96.11	95.4175	95.668	87.22333333	96.002	96.668	88.666	84.668			
	93.34	96.378	98.33	95.692	95.726	89.178	95.998	96.664	88.324	87.9175			
	94.302	96.69	98.44	95.928	96.35	91.836	96.038	96.766	89.19	84.686			
	93.332	96.336	96.082	95.668	95.668	89.668	96.2525	97.085	86.25	83			
	44	41.4	43.2	24.8	19.2	23	20	30.4	21.8	22.4			

La Tabla 5 presenta el desempeño los modelos de redes neuronales convolucionales Compact, AlexNet y ResNet50 evaluados entre imágenes a color y escala de grises, así mismo con diferentes ejes de orientación (XZ, YZ, XY). Al analizar las métricas relevantes para cada condición se destaca que en el caso del modelo compact en el eje XZ, se obtiene una precisión destacada del 98.33%, pero su desempeño disminuye en los ejes YZ (95%) y XY (93.22%), lo que indica que el modelo tiene dificultades para adaptarse a ciertas orientaciones.

En el modelo AlexNet se obtiene un buen desempeño en el eje XZ (100% de exactitud, F1-Score, Precisión y Recall), pero su rendimiento decrece ligeramente en los ejes YZ (96.336%) y XY (95.14%). Se observa que los tiempos de inferencia son rápidos en comparación con los otros dos modelos (máximo 1.8 ms en XY), lo que lo hace eficiente para imágenes en color.

El modelo Resnet 50, cuenta con una precisión en el eje XZ (99.44%) y YZ (97.08%), pero su desempeño en XY cae significativamente (84.67% exactitud). Es notable que los tiempos de inferencia son considerablemente más altos que los de los otros modelos (máximo 107.818 ms en XY).

5.2 Segundo experimento: Conjunto promediado general

El entrenamiento del modelo se realizó considerando el 70% de las imágenes totales (42 imágenes) para entrenar los modelos, permitiendo que aprendieran las características relevantes de las partículas. Un 15% de las imágenes (9 imágenes) se empleó para ajustar los hiperparámetros y evitar el sobreajuste del modelo. El 15% restante de las imágenes (9 imágenes) se usó para evaluar el desempeño final del modelo en datos no vistos durante el entrenamiento.

La ejecución y el entrenamiento se llevaron a cabo en una NVIDIA GeForce RTX 4060 Laptop GPU. Esto permitió acelerar el procesamiento de los datos y las operaciones

matemáticas intensivas, especialmente las relacionadas con las redes neuronales convolucionales.

Se utilizó el optimizador Adam, conocido por combinar las ventajas de los métodos de gradiente descendente estocástico adaptativo.

Hiperparámetros principales:

- **Batch Size:** Se procesaron lotes de 32 imágenes en cada iteración, lo cual es un valor común para balancear el uso de memoria y la estabilidad del gradiente.
- **Iteraciones:** Se realizaron 200 iteraciones totales para cada época.
- **Épocas:** El modelo fue entrenado durante 50 épocas, proporcionando suficientes oportunidades para ajustar los pesos y minimizar la función de pérdida.
- **Learning Rate:** Se utilizó una tasa de aprendizaje de 0.0001, un valor bajo que favorece ajustes precisos de los pesos del modelo.

Tabla 6 Conjunto de datos general promediado

Grises	Grises		Top1-Error	Pérdida	Inferencia por imagen ms	tiempo de procesamiento * ms	Tiempo total* ms	Exactitud %	F1-Score %	Precisión %	Recall %	Mejor época
	Alex net	Compact	5.56	2.78	0.85 ms	1.19 ms	1.29 ms	94.44	94.48	94.59	94.44	28
			3.15E-05	1.05E-03	0.44 ms	0.46 ms	1.65 ms	97.22	97.23	97.28	97.22	50

Colores											
	Resnet 50	Alex net	Compact	Resnet 50							
	3.33	4.44	4.44	4.44							
	2.55E-05	1.02E-05	1.36E-03	9.27E-03							
	134.49 ms	0.95 ms	1 ms	114.93 ms							
	0.9 ms	0.81 ms	0.77 ms	0.61 ms							
	135.39 ms	1.77 ms	1.76 ms	115.54 ms							
	96.67	95.56	95.56	95.56							
	96.68	95.56	95.57	95.58							
	96.74	95.92	95.62	95.7							
	96.67	95.56	95.56	95.56							
	23	17	50	41							

La *Tabla 6* El entrenamiento del modelo se realizó considerando el 70% de las imágenes totales (42 imágenes) para entrenar los modelos, permitiendo que aprendieran las características relevantes de las partículas. Un 15% de las imágenes (9 imágenes) se empleó para ajustar los hiperparámetros y evitar el sobreajuste del modelo. El 15% restante de las imágenes (9 imágenes) se usó para evaluar el desempeño final del modelo en datos no vistos durante el entrenamiento.

La ejecución y el entrenamiento se llevaron a cabo en una NVIDIA GeForce RTX 4060 Laptop GPU. Esto permitió acelerar el procesamiento de los datos y las operaciones

matemáticas intensivas, especialmente las relacionadas con las redes neuronales convolucionales.

Se utilizó el optimizador Adam, conocido por combinar las ventajas de los métodos de gradiente descendente estocástico adaptativo.

Hiperparámetros principales:

- **Batch Size:** Se procesaron lotes de 32 imágenes en cada iteración, lo cual es un valor común para balancear el uso de memoria y la estabilidad del gradiente.
- **Iteraciones:** Se realizaron 200 iteraciones totales para cada época.
- **Épocas:** El modelo fue entrenado durante 50 épocas, proporcionando suficientes oportunidades para ajustar los pesos y minimizar la función de pérdida.
- **Learning Rate:** Se utilizó una tasa de aprendizaje de 0.0001, un valor bajo que favorece ajustes precisos de los pesos del modelo.

Tabla 6 refleja el desempeño de tres arquitecturas de redes neuronales convolucionales (Compact, AlexNet, ResNet50) evaluadas bajo diferentes condiciones de imágenes (en escala de grises y en colores).

Se observó que la arquitectura Compact se destaca en imágenes en escala de grises, alcanzando la mayor precisión del 97.22%, el mejor F1-Score del 97.23%, y manteniendo un tiempo de inferencia por imagen notablemente bajo (1.19 ms). Es el modelo más eficiente para esta condición.

Para imágenes en colores, ResNet50 logra la mejor precisión (96.67%) y F1-Score (96.68%), aunque con tiempos de inferencia significativamente mayores (134.49 ms por imagen).

5.2.1 Comparación de los conjuntos de datos

Comparación de Desempeño

Tabla 7 Comparación de desempeño

Aspecto	Primer Experimento	Segundo Experimento
Precisión (Accuracy)	Compact lideró en grises (97.22%) y ResNet50 en color (96.67%).	AlexNet alcanzó 100% en color (XZ).
Tiempos de Inferencia	Compact mostró la mejor eficiencia, con tiempos mínimos en todas las configuraciones.	Compact sigue liderando en velocidad.
Error Top1	Compact tuvo el menor error en grises (2.78%) y ResNet50 en color (3.33%).	AlexNet y ResNet50 alcanzaron errores de 0 en XZ color.
Consistencia por Modelo	Compact fue estable en grises, mientras que ResNet50 se destacó en color.	AlexNet mostró el mejor balance entre vistas.

Como se observa en la Tabla 7 en la configuración de los experimentos en el experimento se trabajó con imágenes en color y escala de grises evaluando tres modelos: Compact, AlexNet, y ResNet50, con enfoque en la clasificación de partículas con distribuciones específicas para cada vista (XZ, YZ, XY). En el segundo experimento se realiza un análisis más detallado, incluyendo el conjunto de datos en escala de grises y colores.

El segundo experimento confirmó que AlexNet es el modelo más preciso y consistente a lo largo de las iteraciones, especialmente en imágenes a color.

Compact sigue siendo el más eficiente en tiempos de inferencia y presenta un buen rendimiento en escala de grises.

CAPÍTULO VI. CONCLUSIONES Y TRABAJO FUTURO

En el ciclo de desarrollo de esta tesis, se desarrolló y evaluó con éxito una metodología basada en redes neuronales convolucionales profundas, empleando la herramienta Deep Learning Tool para clasificar cascadas extensas de aire (EAS) generadas por rayos cósmicos. Este trabajo permitió la identificación de cascadas utilizando imágenes sintéticas generadas mediante CORSIKA 7.7550, explorando vistas XZ, YZ y XY tanto en color como en escala de grises.

Se entrenaron y evaluaron modelos de redes neuronales convolucionales para clasificar partículas primarias como hierro, protón, muón positivo y rayos gamma.

Se evaluó el rendimiento de diferentes arquitecturas de redes neuronales (ResNet-50, Compact y AlexNet) obteniendo las siguientes conclusiones:

En el experimento utilizando el conjunto de datos por vista promediado, destaca el rendimiento de la red AlexNet en el eje XZ, alcanzando una precisión perfecta del 100% en imágenes a color. Asimismo, en las imágenes a escala de grises, AlexNet demuestra ser el modelo más consistente, superando a Compact y ResNet50 en los ejes XZ y YZ, con precisiones superiores al 96%.

Aunque Compact es el modelo más rápido en todas las configuraciones, con tiempos totales siempre menores a 1.7 ms, AlexNet presenta tiempos ligeramente superiores, pero aún eficientes (máximo 1.8 ms). En contraste, ResNet50 es significativamente más lento, con un tiempo máximo de 107.818 ms en imágenes a color.

En el experimento con el conjunto de datos generales promediados, se observa que las imágenes en escala de grises son ideales para un procesamiento rápido y modelos eficientes, especialmente en contextos con restricciones de hardware en los sitios de observación. Por otro lado, las imágenes a color generan mejores resultados generales, especialmente al utilizar AlexNet y ResNet50, donde la precisión es prioritaria.

En ambos experimentos resalta que AlexNet y Compact se destacan según las necesidades del proyecto, en el caso de AlexNet por su alta precisión y consistencia, y Compact por su rapidez y eficiencia. ResNet50, aunque más lento, es adecuado para tareas

que requieren alta precisión en configuraciones específicas, como los ejes XZ o YZ en imágenes a color.

Se abordó el análisis del perfil vertical de la cascada, un paso inicial hacia la obtención del footprint implementado en la metodología de HAWC. Dado el volumen de imágenes procesadas de manera diaria, la administración eficiente de recursos en los sitios de observación es esencial. En este sentido, la implementación de una clasificación preliminar mediante redes neuronales convolucionales puede optimizar la identificación de datos relevantes y filtrar aquellos considerados como ruido, como las cascadas hadrónicas.

6.1 Limitaciones

- Se requiere aumentar el conjunto de datos expandiendo al análisis a otras partículas primarias para extender el rango de clasificación de las cascadas.
- El tiempo para realizar la simulación es largo por lo que se requiere implementar programación paralela para disminuir los tiempos de las ecuaciones realizadas en la generación de las cascadas en CORSIKA 7.7550.

6.2 Trabajo Futuro

- El campo de estudio en redes neuronales aplicadas a la clasificación de eventos en cascadas atmosféricas se encuentra en constante expansión, por lo que es necesario mejorar y automatizar los procesos de clasificación con herramientas simples que permitan implementar estas metodologías.
- La versión de CORSIKA 8 actualmente se encuentra en desarrollo, la implementación en futuros trabajos de esta versión podría contener correcciones en los modelos implementados, así como una mejor adaptación para lenguajes de programación como Python.
- Se considera importante los análisis dinámicos en vistas tridimensionales y en la línea temporal de las cascadas, lo que podría revelar detalles importantes, como procesos de decaimiento y extinción de partículas.
- Integrar estas metodologías en sistemas de detección en tiempo real ayudaría a mejorar los porcentajes al identificar cascadas relevantes en experimentos como HAWC, Pierre Auger o el futuro SWGO.

REFERENCIAS BIBLIOGRÁFICAS

- (CTAO), C. T. (n.d.). Cherenkov Effect. Obtenido de <https://www.ctao.org/es/emission-to-discovery/science/how-ctao-works/>
- Arteaga Velázquez, J. (2020). HAWC, OBSERVATORIO ASTROFÍSICO DE RAYOS GAMMA. *Revista Saber Más*, págs. https://www.sabermas.umich.mx/archivo/articulos/227-numero-26/408-hawc-observatorio-astrofisico-de-rayos-gamma.html?utm_source=chatgpt.com.
- Bleicher, M., Zabrodin, E., Spieles, C., Bass, S. A., Ernst, C., Soff, S., ... & Stöcker, H. . (1999). Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model. *Journal of Physics G: Nuclear and Particle Physics*, 25(9), 1859. doi:10.1088/0954-3899/25/9/308
- Cherenkov Telescope Array Observatory (CTAO). (s.f.). *How CTAO works*. . Obtenido de <https://www.ctao.org/es/emission-to-discovery/science/how-ctao-works/>
- Drew, R. D. (2021). Deep unsupervised domain adaptation for gamma-hadron separation . *Technische Universität Dortmund*.
- Engel, R., Heck, D., & Pierog, T. (2011). xtensive air showers and hadronic interactions at high energy. *Annual Review of Nuclear and Particle Science*, 61, 467–489. Obtenido de <https://doi.org/10.1146/annurev.nucl.012809.104544>
- Ferrari, A., Sala, P. R., Fassò, A., & Ranft, J. (2024). *Fluka: A multi-particle transport code (Program version 2024)*. Obtenido de <http://www.fluka.org>

Glombitza, J. . (2017). *A deep learning-based reconstruction of air showers at the Pierre Auger Observatory (Tesis de maestría, RWTH Aachen University, Faculty of Mathematics, Computer Science and Natural Sciences)*.

Hampel-Arias, Z., & Westerhoff, S. (2015). Gamma hadron separation using pairwise compactness method with HAWC. . *34th International Cosmic Ray Conference* . , 1-8. Obtenido de arXiv:1508.04047v2

Hampel-Arias, Z., & Westerhoff, S. (2015)). (. Gamma hadron separation using pairwise compactness method with HAWC. *34th International Cosmic Ray Conference*, 1-8. Obtenido de arXiv:1508.04047v2

Heck, D., & Pierog, T. (2023). *Extensive Air Shower Simulation with CORSIKA: A User's Guide (Version 7.7500 from April 14, 2023)*. Karlsruher Institut für Technologie (KIT), Institut für Astroteilchenphysik.

Luzquiños Saavedra, D. A. (2024). *Gamma-Hadron separation using the temporal distribution of particle cascades at TeV energies in the SWGO experiment (Tesis de licenciatura, Pontificia Universidad Católica del Perú, Facultad de Ciencias e Ingeniería)*. Obtenido de https://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/27943/LUZQUIÑOS_SAAVEDRA_DAVID_%20GAMMA_HADRON_SEPARATION.pdf?sequence=1&isAllowed=y

MVTec Software GmbH. (2024). *MVTec Deep Learning Tool* . Obtenido de https://software.mvtec.com/acton/attachment/43208/f-df8afe18-f3c6-4622-b429-498f353d6252/1/-/-/-/2024.05.14_leaflet_deep_learning_tool_rgb_en.pdf

- Ohm, S. (2010). Advanced hadron separation technique and application to particular γ -ray sources with H.E.S.S. . *University of Heidelberg*.
- Sharma, M., Nayak, J., Koul, M. K., Bose, S., & Mitra, A. (2014). Gamma/hadron segregation for a ground based imaging. *Astrophysical Sciences Division, Bhabha Atomic Research Centre., 14(11)*, 1491–1503. doi:10.1088/1674–4527/14/11/012
- Zuñiga Reyes, A. (2017). *Estudio de muones para separación Gamma/Hadrones en el observatorio HAWC (Tesis de maestría)*. Universidad Nacional Autónoma de México.

ACRÓNIMOS

ANN	Redes Neuronales Artificiales
CNNs	Convolutional Neural Network
CR's	Cosmic Rays
DISC	Análisis Discriminante Lineal
EAS	Extensive Air Shower
HAWC	High-Altitude Water Cherenkov
NB	Clasificador Naive Bayes
RF	Random Forest
SVM	Máquinas de Soporte Vectorial