

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

DERLEI JURANDIR DA SILVA

**CLASSIFICAÇÃO E RECONSTRUÇÃO DE CHUVEIROS ATMOSFÉRICOS VIA
DEEP LEARNING ESTEREOSCÓPICO MULTITAREFA PARA OS
TELESCÓPIOS DE GRANDE PORTE DO CTAO**

CURITIBA

2026

DERLEI JURANDIR DA SILVA

**CLASSIFICAÇÃO E RECONSTRUÇÃO DE CHUVEIROS ATMOSFÉRICOS VIA
DEEP LEARNING ESTEREOSCÓPICO MULTITAREFA PARA OS
TELESCÓPIOS DE GRANDE PORTE DO CTAO**

**Classification and Stereoscopic Reconstruction of Air Showers via
Multi-task Deep Learning for CTAO Large-Sized Telescopes**

Dissertação apresentada como requisito para obtenção do título de Mestre em Física e Astronomia do Programa de Pós Graduação em Física e Astronomia da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Alexandre José Tuoto
Silveira Mello

CURITIBA

2026



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



DERLEI JURANDIR DA SILVA

**CLASSIFICAÇÃO E RECONSTRUÇÃO DE CHUVEIROS ATMOSFÉRICOS VIA DEEP LEARNING
ESTEREOSCÓPICO MULTITAREFA PARA OS TELESCÓPIOS DE GRANDE PORTE DO CTAO**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Física da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Física.

Data de aprovação: 16 de Março de 2026

Dr. Alexandre Jose Tuoto Silveira Mello, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Carlos Jose Toderer Peixoto, Doutorado - Universidade de São Paulo

Dra. Rita De Cassia Dos Anjos, Doutorado - Universidade Federal do Paraná (Ufpr)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 16/03/2026.

Em algum lugar na vastidão do espaço-tempo,
há uma jovem mente curiosa que aguarda
ansiosa e com brilho nos olhos o próximo
episódio de Espaçonave Terra.

Este trabalho é dedicado para a mãe, o pai e o
irmão dela.

AGRADECIMENTOS

À minha família - meu pai, minha mãe e meu irmão - por serem constantes em um Universo de variáveis, sempre me apoiando e incentivando desde os meus primeiros e titubeantes passos às margens do oceano cósmico.

Agradeço aos docentes e à coordenação do PPGFA, especialmente ao meu orientador, Prof. Dr. Alexandre José Tuoto Silveira Mello, pela paciência e confiança investidas em mim ao longo destes anos.

Ao CAUTEC, sobretudo às amigadas que dele surgiram. Agradeço pelas noites de observação e jogatinas de RPG, pela companhia diária, pelos foguetes lançados e por compartilharem a esperança de um mundo melhor levando a Astronomia para fora da Universidade.

Agradeço, em particular, a quem esteve presente e acompanhou de perto o desenvolvimento deste trabalho. Aos amigos e alunos que cansaram de ouvir o quão incrível é a astronomia de raios gama.

Esta pesquisa utilizou os Modelos de Simulação de Telescópios do CTAO, fornecidos pela Organização Central e pelo Consórcio do CTAO.

Este trabalho foi realizado com o apoio financeiro da Fundação Araucária e da Secretaria de Estado da Ciência, Tecnologia e Ensino Superior por meio do PDI 346/2024 - NAPI Fenômenos Extremos do Universo. Parte do trabalho foi desenvolvido utilizando a infraestrutura computacional do PPGFA financiada pela Fundação Araucária (Computador ATLAS).

"terras licet [...] et undas obstruat:

et caelum certe patet;

ibimus illac:

omnia possideat, non possidet aera..."

OVIDI NASONIS METAMORPHOSEON,
LIBRI VIII.

RESUMO

O *Cherenkov Telescope Array Observatory (CTAO)* representa a próxima geração de observatórios de raios gama terrestres, projetado para melhorar a sensibilidade em uma ordem de grandeza em relação aos instrumentos atuais. Um dos principais desafios científicos do CTAO reside na faixa de baixa energia (20-150 GeV), coberta pelos *Large-Sized Telescopes (LSTs)*, onde a escassez de fótons Cherenkov e a contaminação pelo fundo hadrônico limitam as técnicas de reconstrução tradicionais baseadas em parametrização de Hillas. Esta dissertação propõe e valida uma rede neural convolucional profunda projetada para a reconstrução estereoscópica completa de eventos nos LSTs. A arquitetura opera sob o paradigma de aprendizado multitarefa (*Multi-Task Learning*), realizando simultaneamente a classificação gama-hádron e a regressão de energia e direção de chegada. Foi implementada uma estratégia de *Curriculum Learning*, evoluindo a rede desde o aprendizado morfológico básico até a correlação geométrica complexa. Utilizando simulações de Monte Carlo (CORSIKA+sim_telarray), demonstramos que a fusão de dados estereoscópicos aliada a mecanismos de atenção permite atingir, no regime de operação estereoscópico total, uma resolução de energia abaixo de 20% e uma resolução angular θ_{68} na ordem de $0,2^\circ$. A classificação gama-hádron apresentou uma Área Sob a Curva (AUC) superior a 0,97. Mantendo alta fidelidade energética e capacidade de classificação também em fontes difusas, essenciais para o mapeamento de estruturas extensas e realização de *surveys*. Consolidando a rede proposta como uma ferramenta promissora para a maximização do potencial científico da próxima geração de telescópios Cherenkov.

Palavras-chave: aprendizado multitarefa; astronomia de raios gama; ctao; estereoscopia.

ABSTRACT

The Cherenkov Telescope Array Observatory (CTAO) represents the next generation of ground-based gamma-ray observatories, designed to improve sensitivity by an order of magnitude over current instruments. A major scientific challenge for CTAO lies in the low-energy range (20-150 GeV), covered by the Large-Sized Telescopes (LSTs), where the scarcity of Cherenkov photons and hadronic background contamination limit the traditional reconstructions methods based on Hillas parametrization. This work proposes and validates a deep convolutional siamese neural network designed for full stereoscopic event reconstruction in LSTs. The architecture operates under the Multi-Task Learning paradigm, simultaneously performing gamma-hadron classification and energy and arrival direction regression. A Curriculum Learning strategy was implemented, evolving the neural network from basic morphological learning to complex geometric correlation. Using Monte Carlo simulations (CORSIKA+sim_telarray), we demonstrate that stereoscopic fusion combined with attention mechanisms permit achieving, in a total stereoscopic operation, an energy resolution under 20% and an angular resolution θ_{68} of $0,2^\circ$. The gamma-hadron classification yielded an Area Under the Curve (AUC) exceeding 0,97. The network maintains high energy fidelity and classification capacity for diffuse sources, which are essential for mapping extended structures and conducting surveys. These results consolidate the proposed network as a promising tool for maximizing the scientific potential of the next generation of Cherenkov telescopes.

Keywords: ctao; gamma-ray astronomy; multitask learning; stereoscopy.

LISTA DE FIGURAS

Figura 1 – Espectro eletromagnético, o espectro visível ocupa apenas uma estreita faixa	23
Figura 2 – Esboço técnico do eletroscópio de Coulomb	24
Figura 3 – Victor Hess em um de seus vôos, entre 1911 e 1912	25
Figura 4 – Chapas fotográficas da descoberta do Méson π	27
Figura 5 – Fusão de duas estrelas de nêutrons observada em diferentes sinais, considerado um dos eventos de maior destaque da astronomia multi-mensageira	28
Figura 6 – Trajetória e interação das partículas no espaço	28
Figura 7 – Critério de Hillas, ilustrando diferentes classes de aceleradores: O eixo vertical representa a intensidade do campo magnético enquanto o eixo horizontal representa o tamanho do objeto, as linhas sólidas representam o limite de energia na qual a partícula "escapa" dos limites do objeto, enquanto as linhas tracejadas indicam as regiões do Joelho e Tornezel no espectro.	31
Figura 8 – Mecanismo de Fermi de segunda ordem	32
Figura 9 – Mecanismo de Fermi de primeira ordem	33
Figura 10 – Aceleração por múltiplas reflexões entre ondas de choque	34
Figura 11 – Espectro de emissão Gama: Bremsstrahlung não é representado por ser subdominante no intervalo	35
Figura 12 – Produção de radiação síncrotron pela deflexão de partículas carregadas em um campo magnético	36
Figura 13 – Emissão de raio gama por bremsstrahlung	37
Figura 14 – Geração de raios gama via bremsstrahlung pela colisão entre nuvens	38
Figura 15 – Ilustração do mecanismo de emissão de raios gama via espalhamento Compton inverso	39
Figura 16 – Produção de pión neutro por interação próton-núcleo e posterior decaimento em um par fóton-fóton	40
Figura 17 – Fluxo observado de raios cósmicos de alta energia: eventos de ultra-alta energia são raros	41

Figura 18 – Estrutura típica de um remanescente de supernova - Nebulosa do Caranguejo	42
Figura 19 – Pulsar na região central da Nebulosa do Caranguejo	43
Figura 20 – Formação do disco de acreção em um sistema binário com objeto compacto	43
Figura 21 – Núcleo Ativo de Galáxia (AGN), um buraco negro central é responsável pela alta luminosidade desses objetos	44
Figura 22 – Modelo de bola de fogo para explosões de raios gama	44
Figura 23 – Desenvolvimento longitudinal e lateral de EAS iniciado por um raio gama de 1 TeV	46
Figura 24 – Desenvolvimento longitudinal e lateral de EAS iniciado por um próton de 1 TeV	46
Figura 25 – a) Modelo de Heitler para cascatas eletromagnéticas e b) Modelo de Heitler-Matthews para cascatas hadrônicas	48
Figura 26 – Diferentes técnicas utilizadas para a detecção de EAS	48
Figura 27 – Polarização do meio e emissão de radiação Cherenkov	49
Figura 28 – Ângulo de emissão de radiação Cherenkov em diferentes altitudes na atmosfera	50
Figura 29 – Atualmente existem mais de 320 fontes catalogadas, entre galáticas, extragaláticas e desconhecidas. Informações obtidas do site TeVCat em março de 2026	51
Figura 30 – Esquema simplificado do sistema de detecção de um telescópio Cherenkov.	52
Figura 31 – Distribuição lateral, temporal e angular dos fotoelétrons capturados pela câmera em um IACT	53
Figura 32 – Sensibilidade projetada para o CTAO com 50h de observação em comparação com outros instrumentos	54
Figura 33 – Comparação em escala mostrando a altura das três classes de telescópios projetadas para o CTAO	55
Figura 34 – Cluster de PMTs, contendo 21 pixels cada, acoplados a um sistema eletrônico de leitura ultra-rápida	55

Figura 35 – Distribuição geográfica dos telescópios da configuração alfa no sítio CTAO Norte, além dos 4 LSTs e dos 9 MSTs, a imagem também mostra os 2 telescópios MAGIC	56
Figura 36 – Distribuição geográfica dos telescópios da configuração alfa no sítio CTAO Sul, além dos 14 MSTs e dos 37 SSTs, a imagem também mostra os locais de fundação para futura expansão com 3 SSTs e 4 LSTs.	57
Figura 37 – Parametrização de Hillas de uma imagem obtida por IACTs	58
Figura 38 – Simulações IACT com CORSIKA e sim_telarray	59
Figura 39 – Imagens obtidas em simulação de Monte Carlo	61
Figura 40 – Fluxo de dados do CTAO após observações, os dados brutos (DL0) passam por etapas de processamento antes de serem disponibilizados ao público (DL3)	61
Figura 41 – Desenho esquemático de um perceptron, contendo entradas que são multiplicadas por um peso e resultam em uma saída	62
Figura 42 – Esquemático de um MLP, camadas mais profundas extraem informações mais complexas	63
Figura 43 – Diferentes funções de ativação utilizadas em redes neurais artificiais	64
Figura 44 – Função Rectified Linear Unit - ReLU	64
Figura 45 – Função Sigmoide	65
Figura 46 – Algoritmos de otimização de gradiente. a) Gradiente descendente ($\alpha = 0.05$) b) Gradiente descendente ($\alpha = 1$) c) Gradientes normalizados ($\alpha = 0.05$) d) ADAM ($\alpha = 0.05, \beta = 0.9, \gamma = 0.99$)	68
Figura 47 – Operação de convolução do kernel $h(x,y)$ com a imagem $f(x,y)$, resultando em $g(x,y)$. Os pixels em azul representam os pixels que deram origem ao pixel em verde	69
Figura 48 – a) Arquitetura de uma CNN LeNet-5, utilizada em tarefas de reconhecimento de dígitos. b) Visualização das features aprendidas pela rede.	70
Figura 49 – Esquemático de uma rede multitarefa, o <i>backbone</i> alimenta sub-redes específicas para cada tarefa	71
Figura 50 – Matriz de confusão	71

Figura 51 – Curva ROC. A diagonal mostra o desempenho de um classificador aleatório. Três exemplos de classificadores (azul, laranja e verde) são apresentados.	73
Figura 52 – Imagens de um EAS iniciado por gama (~ 60 GeV) a 40°	77
Figura 53 – Imagens de um EAS iniciado por próton (~ 60 GeV) a 40°	77
Figura 54 – Distribuição energética dos eventos de treino	78
Figura 55 – Distribuição energética dos eventos de validação	78
Figura 56 – Representação do método de mapeamento por vizinho mais próximo	80
Figura 57 – Bloco STEM, responsável por extrair as primeiras características e preparar a imagem para a rede	82
Figura 58 – Blocos convolucionais residuais	82
Figura 59 – Bloco de <i>Squeeze-and-Excitation</i>	82
Figura 60 – Mecanismo de <i>multi-head self-attention</i>	83
Figura 61 – Esquemático do bloco denso utilizado no <i>neck</i>	83
Figura 62 – As 4 heads, compostas por camadas densas	84
Figura 63 – Representação simplificada da rede neural	85
Figura 64 – Valores da Loss Global (Incerteza Homocedástica) para o treino e validação no primeiro estágio de treinamento	89
Figura 65 – Curva ROC, o modelo alcança uma AUC de 95% ainda nas fases iniciais do treinamento	90
Figura 66 – Evolução da função de perda para as tarefas de regressão de Energia e Direção durante o primeiro estágio de treinamento.	90
Figura 67 – Loss da direção, mostrando que a rede atinge a convergência da geometria estéreo	91
Figura 68 – Função de perda das tarefas de regressão durante o aprendizado de energia	91
Figura 69 – Função de perda das tarefas de regressão da energia e direção ao final do treinamento, com regime de exploração da taxa de aprendizagem	92
Figura 70 – Curva ROC para o cenário global, alcançando AUC de 0,934	92
Figura 71 – Limiar de decisão (Gamma _{ness}) em azul, prótons em cinza	93
Figura 72 – Matriz de confusão para o cenário global	94
Figura 73 – Matriz de migração de energia para o cenário global	94

Figura 74 – Resolução Angular (Região de 68% de contenção) do cenário global . .	95
Figura 75 – Matriz de confusão stereo	96
Figura 76 – Resolução Angular em um cenário de estereoscopia para 2, 3 e 4 teles- cópios	97
Figura 77 – Matriz de migração de energia em diferentes cenários de multiplicidade	97
Figura 78 – Desempenho da classificação (AUC) por energia em cenários com es- tereoscopia	98
Figura 79 – Resolução Angular θ_{68} por níveis de energia em cenários com estere- oscopia	99

LISTA DE TABELAS

Tabela 1 – Comparação de Desempenho entre Cenários de Multiplicidade	96
Tabela 2 – Desempenho por ângulo zenital e faixa de energia em observação este- reoscópica ($\geq 2 Tels$).	99
Tabela 3 – Comparativo de desempenho entre as componentes pontuais e difusas (4 Tels, 40°, 50-150 GeV).	100
Tabela 4 – Especificações da infraestrutura computacional e ambiente de software utilizados no treinamento da rede neural.	110

LISTA DE ABREVIATURAS E SIGLAS

Siglas

ADAM	<i>ADaptive Moment estimation</i>
AGN	<i>Active Galactic Nuclei</i>
AUC	<i>Area Under the Curve</i>
BCE	<i>Binary Cross-Entropy</i>
CMB	<i>Cosmic Microwave Background</i>
CNN	<i>Convolutional Neural Networks</i>
CORSIKA	<i>COsmic Ray Simulations for KAscade</i>
CTAO	<i>Cherenkov Telescope Array Observatory</i>
DL	<i>Deep Learning</i>
EAS	<i>Extensive Air Shower</i>
EBL	<i>Extragalactical Background Light</i>
EHE	<i>Extremely-High Energy</i>
GRB	<i>Gamma-Ray Burst</i>
HE	<i>High Energy</i>
HESS	<i>High Energy Stereoscopic System</i>
IACT	<i>Imaging Atmospheric Cherenkov Technique</i>
LE	<i>Low Energy</i>
LST	<i>Large-Sized Telescope</i>
MAGIC	<i>Major Atmospheric Gamma Imaging Cherenkov</i>
ML	<i>Machine Learning</i>
MLP	<i>Multi Layer Perceptron</i>
MSE	<i>Mean Squared Error</i>

MST	<i>Medium-Sized Telescope</i>
MTL	<i>Multi Task Learning</i>
PMT	<i>Photo Multiplier Tube</i>
PWNe	<i>Pulsar Wind Nebula</i>
RCS	<i>Raios C3smicos</i>
ReLU	<i>Rectified Linear Unit</i>
RNA	<i>Rede Neural Artificial</i>
ROC	<i>Receiver Operating Characteristic</i>
SE	<i>Squeeze-and-Excitation</i>
SiPM	<i>Silicon Photomultiplier</i>
SNR	<i>Supernova Remnant</i>
SST	<i>Small-Sized Telescope</i>
STL	<i>Single Task Learning</i>
UHE	<i>Ultra-High Energy</i>
VERITAS	<i>Very Energetic Radiation Imaging Telescope Array System</i>
VHE	<i>Very-High Energy</i>

SUMÁRIO

1	INTRODUÇÃO	20
1.1	Astronomia de Raios Gama	20
1.2	Chuveiro Atmosférico Extenso	20
1.3	Imageamento Cherenkov Atmosférico	21
1.4	Desafios da Reconstrução de Eventos	21
1.5	Inteligência Artificial aplicada à Astrofísica	21
1.6	Objetivos	22
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	Mensageiros Cósmicos	23
2.2	Raios Gama	27
2.2.1	Produção e Propagação	29
2.2.1.1	Mecanismos de Aceleração	30
2.2.1.1.1	<i>Mecanismo de Fermi de Segunda Ordem</i>	32
2.2.1.1.2	<i>Mecanismo de Fermi de Primeira Ordem</i>	33
2.2.1.2	Geração de Raios Gama de Altas Energias	35
2.2.1.2.1	<i>Síncrotron</i>	36
2.2.1.2.2	<i>Bremsstrahlung</i>	36
2.2.1.2.3	<i>Espalhamento Compton Inverso</i>	38
2.2.1.2.4	<i>Decaimento de mésons neutros</i>	40
2.2.2	Fontes Astrofísicas	41
2.2.2.1	Propagação	45
2.3	Chuveiros Atmosféricos Extensos	45
2.3.1	Chuveiros eletromagnéticos	46
2.3.2	Chuveiros hadrônicos	47
2.3.3	Emissão de Radiação Cherenkov em EAS	49
2.4	Imageamento Cherenkov Atmosférico	50
2.4.1	Princípios de funcionamento	51
2.4.2	Cherenkov Telescope Array Observatory	52
2.4.2.1	CTAO - Norte	54
2.4.2.2	CTAO - Sul	55

2.5	Métodos Tradicionais de Reconstrução	56
2.5.0.1	Abordagem paramétrica - Parâmetros de Hillas	56
2.5.0.2	Abordagem Baseada em Modelos	58
2.6	Simulações Monte Carlo	58
2.6.1	CORSIKA	59
2.6.2	sim_telarray	60
2.7	Redes Neurais Artificiais e Aprendizado Profundo	61
2.7.1	Fundamentos de Redes Neurais	62
2.7.1.1	Função de ativação	63
2.7.1.1.1	<i>Rectified Linear Unit (ReLU)</i>	64
2.7.1.1.2	<i>Sigmoide</i>	65
2.7.2	Treinamento	65
2.7.2.0.1	<i>Função de perda</i>	65
2.7.2.0.2	<i>Retropropagação ou otimização gradiente</i>	67
2.7.3	Aprendizado Curricular	68
2.7.4	Redes Neurais Convolucionais	69
2.7.5	Aprendizado Multitarefa	70
2.7.5.1	Métricas de avaliação	71
3	MÉTODOS	74
3.1	Simulação de eventos	74
3.1.1	CORSIKA	74
3.1.2	Compilação	74
3.1.3	Arquivos de <i>Input</i>	75
3.1.4	O pacote sim_telarray	76
3.2	Conjunto de dados	76
3.2.1	Obtenção dos dados e Pré-Processamento	79
3.3	Arquitetura da Rede Neural	81
3.3.1	<i>Backbone</i> siamês	81
3.3.2	Gargalo (<i>Neck</i>)	83
3.3.3	Decodificadores (<i>Heads</i>)	83
3.3.4	Pooling Híbrido	84
3.3.4.1	Global Average Pooling	85

3.3.4.2	Global Sum Pooling	85
3.4	Detalhes de Implementação	85
3.5	Métricas de Avaliação	86
3.5.1	Métricas de Classificação (Separação Gama/Hádron)	86
3.5.1.1	Curva ROC e AUC	86
3.5.1.2	Matriz de Confusão	87
3.5.1.3	Precisão, Recall e F1-Score	87
3.5.1.4	Acurácia	87
3.5.2	Métricas de Reconstrução de Energia	87
3.5.2.1	Viés de Energia	87
3.5.2.2	Resolução Energética	88
3.5.3	Matriz de Migração	88
3.5.4	Métricas de Reconstrução da Direção	88
4	RESULTADOS E DISCUSSÃO	89
4.1	Análise da Evolução do Treinamento e Convergência do modelo	89
4.2	Classificação Gama-Hádron: Global	92
4.3	Matriz confusão e limiar de decisão: Global	93
4.4	Reconstrução de Energia: Global	94
4.5	Reconstrução da Direção: Global	95
4.6	Análise de Dependências: Multiplicidade, Energia e Ângulo Zenital	95
4.6.1	Multiplicidade	96
4.6.2	Energia	97
4.6.3	Ângulo Zenital	98
4.6.4	Gama pontual vs Gama Difuso	100
5	CONCLUSÃO	101
5.1	Principais Resultados	101
5.2	Contribuições para a área	102
5.3	Trabalhos Futuros	102
	REFERÊNCIAS	103

APÊNDICES	108
APÊNDICE A – INFRAESTRUTURA COMPUTACIONAL	110

1 INTRODUÇÃO

A astrofísica de partículas é a Ciência que investiga os fenômenos mais energéticos do Universo, utilizando partículas cósmicas como mensageiros. A observação do Cosmos não se limita mais apenas a luz visível, como foi durante milênios. Hoje vivemos na era da astronomia multimessageira, que combina a detecção de ondas eletromagnéticas, neutrinos, raios cósmicos e ondas gravitacionais para compor uma imagem completa dos objetos astrofísicos (SPURIO, 2018; LONGAIR, 2011; GRUPEN, 2020).

1.1 Astronomia de Raios Gama

A astronomia de raios gama estuda a forma mais energética de radiação eletromagnética, cobrindo uma extensa faixa de energia que se estende pelo espectro eletromagnético a partir da escala de MeV. Ao contrário dos raios cósmicos carregados, que são desviados por campos magnéticos e perdem informação direcional de sua origem, os raios gama são partículas eletricamente neutras que viajam em linha reta, apontando, portanto, diretamente para as suas fontes (SPURIO, 2018).

As fontes de raios gama de muito alta energia estão associadas a ambientes astrofísicos extremos onde ocorre a aceleração de partículas a velocidades relativísticas, como, por exemplo: remanescentes de supernovas (*Supernova Remnant* (SNR)), nebulosas de vento de pulsar (*Pulsar Wind Nebula* (PWNe)), Núcleos Ativos de Galáxias (*Active Galactic Nuclei* (AGN)), etc (ACHARYA *et al.*, 2018). Permitindo estudar os mecanismos físicos que atuam em ambientes extremos (LONGAIR, 2011).

No entanto, a observação direta dessas partículas de altas energias a partir da superfície terrestre geralmente é impossibilitada pela sua interação com a atmosfera, exigindo métodos de detecção indireta. (GRUPEN, 2020).

1.2 Chuveiro Atmosférico Extenso

Quando um raio gama ou um raio cósmico de alta energia colide com um núcleo na atmosfera terrestre, ele inicia uma cascata de partículas secundárias em um evento conhecido como Chuveiro Atmosférico Extenso (*Extensive Air Shower* (EAS)) (ARBELETICHE, 2021). O desenvolvimento de um EAS depende da partícula primária: um chuveiro eletromagnético, iniciado por um fóton ou elétron, é dominado por processos de produção de pares (e^-e^+) e *Bremsstrahlung*. O número de partículas na cascata cresce exponencialmente até atingir um máximo e, em seguida, atenua-se à medida que a energia média das partículas cai abaixo de uma energia crítica. Por outro lado, a vasta maioria dos chuveiros é iniciada por hádrons, como

prótons e núcleos mais pesados, que além da componente eletromagnética, também apresentam componente hadrônica e muônica (SANT'ANNA JUNIOR, 2022).

1.3 Imageamento Cherenkov Atmosférico

Embora as partículas da cascata sejam absorvidas pela atmosfera, muitas delas viajam a velocidades superiores à velocidade da luz no ar, emitindo radiação Cherenkov. Essa radiação é emitida em um cone ao redor da trajetória da partícula e chega ao solo como um breve *flash* de luz mais intenso entre 380 nm e 480 nm, com duração de alguns nanossegundos.

A técnica de Imageamento Cherenkov Atmosférico (*Imaging Atmospheric Cherenkov Technique* (IACT)) utiliza grandes espelhos para coletar essa luz e focá-la em um sensor. A imagem resultante reflete a distribuição angular e espacial do chuveiro. A observação moderna, em observatórios como o MAGIC e o HESS, é realizada em modo estereoscópico, utilizando múltiplos telescópios para observar o mesmo chuveiro de diferentes ângulos, o que melhora a resolução angular, a resolução em energia e a capacidade de rejeição de ruído de fundo (SOUTO, 2018). Atualmente o futuro da área reside no CTAO, que melhorará a sensibilidade em uma ordem de grandeza e expandirá a cobertura de energia de 20 GeV a 300 TeV.

1.4 Desafios da Reconstrução de Eventos

O principal desafio na análise de dados de IACTs é a separação entre sinal e ruído. Para cada raio gama detectado de uma fonte, existem milhares de eventos de fundo causados por raios cósmicos hadrônicos. Tradicionalmente, a reconstrução e classificação de eventos baseia-se na parametrização das imagens usando parâmetros propostos por HILLAS (1985). Com o advento da computação, a classificação e reconstrução desses chuveiros passa a ser híbrida, assim métodos de aprendizado de máquina supervisionado, como *Random Forest*, são treinados com esses parâmetros para separar raios gama de hádrons. Essa parametrização pode descartar informações sutis contidas na imagem, e quando se trata de baixas energias, raramente esses parâmetros ficam evidentes, portanto o método pode inserir limitações na sensibilidade do instrumento (KRYUKOV; DEMICHEV; ILYIN, 2024).

1.5 Inteligência Artificial aplicada à Astrofísica

Nos últimos anos, as técnicas de Aprendizado Profundo (*Deep Learning* (DL)), especificamente Redes Neurais Convolucionais (*Convolutional Neural Networks* (CNN)), têm sido implementadas como uma alternativa promissora aos métodos tradicionais. As CNNs são capazes de extrair características diretamente das imagens, sem a necessidade de uma parametrização prévia.

A aplicação de CNNs a dados de IACTs apresenta desafios específicos, como o fato de muitas câmeras utilizarem pixels hexagonais, o que requer adaptações nas operações de convolução padrão ou pré-processamento das imagens para matrizes quadradas. Estudos recentes demonstraram que as CNNs podem superar os métodos clássicos na classificação de eventos (SILVA; MELLO, 2025; JACQUEMONT *et al.*, 2021; NIETO *et al.*, 2019).

1.6 Objetivos

Este trabalho tem como objetivo principal desenvolver e avaliar uma arquitetura de Rede Neural Multitarefa para a classificação e reconstrução de eventos de raios gama no contexto dos *Large Sized Telescopes (LSTs)* do *Cherenkov Telescope Array Observatory (CTAO)*.

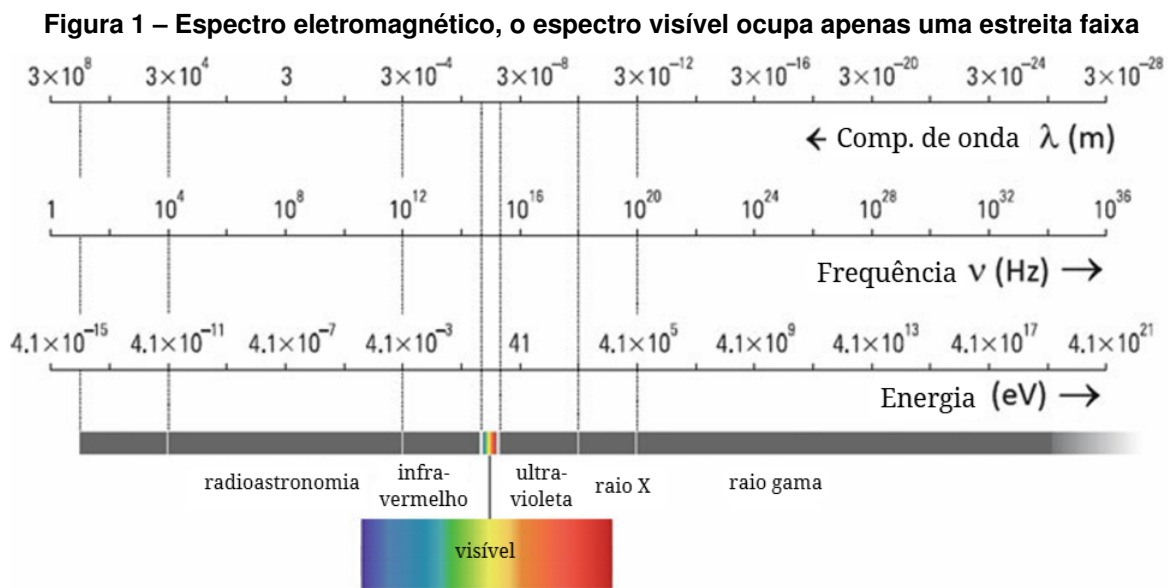
Os objetivos específicos são:

1. Gerar um conjunto de dados robusto de chuveis atmosféricos iniciados por raios gama e prótons (fundo hadrônico) utilizando os softwares de simulação CORSIKA (HECK *et al.*, 1998) e sim_telarray (BERNLÖHR, 2008), garantindo controle total sobre os parâmetros de simulação.
2. Implementar uma cadeia de pré-processamento para converter as imagens de pixels hexagonais das câmeras dos telescópios para um formato de matriz quadrada compatível com arquiteturas de CNN padrão.
3. Treinar e validar um modelo de rede neural multitarefa, capaz de realizar tanto a classificação binária (gama/hádron), quanto a reconstrução de energia e direção, explorando a visão estereoscópica do arranjo.
4. Avaliar o desempenho do modelo na faixa de energia de máxima eficiência do LST.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Mensageiros Cósmicos

Durante a maior parte da história humana, a compreensão acerca do Cosmos restringiu-se às informações obtidas por observações no espectro visível, uma estreita faixa de radiação eletromagnética que compreende fótons com comprimento entre ~ 376 nm (violeta) a ~ 780 nm (vermelho), ou, em termos usuais adotados na astrofísica de partículas: 1,6 eV (vermelho) e 3,3 eV (violeta). A Figura 1 ilustra o espectro eletromagnético, destacando a faixa do espectro visível. (LONGAIR, 2011; SPURIO, 2018).

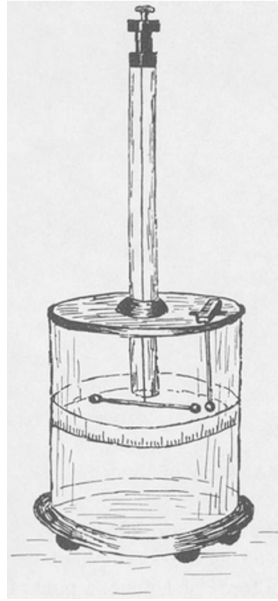


Fonte: Spurio (2018).

Foi somente no início do século XX que a humanidade passou a olhar para o que não era visível. Embora a literatura costumeiramente assinale os experimentos com balão de HESS (1912) como marco inicial da astrofísica de partículas, Dorman e Dorman (2014) mostram que "A história da descoberta dos raios cósmicos pode ser considerada um exemplo das consequências significativas que podem surgir a partir da investigação rigorosa de um efeito desconhecido e, aparentemente, insignificante." (DORMAN; DORMAN, 2014, p. 2), nesse caso, a ionização constante do ar.

Em 1785, Charles Augustin de Coulomb percebeu que seu eletroscópio (Figura 2) perdia gradualmente a carga, apesar de todas suas preocupações para garantir que ele estivesse isolado. Inicialmente, Coulomb atribuiu esse fenômeno ao contato de partículas de poeira e vapor d'água presentes no ar, que transportariam a carga para longe do corpo eletrizado. A próxima descoberta significativa veio com William Crookes, que, em 1879, observou que esse efeito diminuía significativamente sob pressão reduzida, indicando que a ionização do ar era a causa direta da dispersão das cargas elétricas.

Figura 2 – Esboço técnico do eletroscópio de Coulomb



Fonte: Gamow (1981).

No final do século XX, impulsionados pela descoberta dos raios X, por Wilhelm Conrad Röntgen, e da radioatividade natural, por Antoine Henri Becquerel, físicos como J.J. Thomson demonstraram que a condutividade em gases aumentava drasticamente sob influência de radiações ionizantes. Independentemente, Hans Geitel e Julius Elster, na Alemanha, e Charles Thomson Rees Wilson na Inglaterra, confirmaram em 1900 que a descarga ocorria mesmo em recipientes hermeticamente fechados e livres de poeira, sugerindo a presença de uma radiação ionizante onipresente. Wilson inclusive calculou que cerca de 20 íons eram formados por segundo em cada centímetro cúbico de ar.

Nesse período, embora Marie Curie^{1 2} e o próprio Wilson³ tenham levantado a hipótese de que essa ionização fosse extraterrestre, o consenso era de que essa ionização residual aparecia devido a radioatividade natural da crosta terrestre. De fato, Dorman e Dorman (2014) esclarece que os experimentos iniciais pareciam corroborar essa visão: medições em cavernas e túneis mostravam níveis de radiação consistentes ou elevados, levando a falsos negativos quanto a uma origem externa.

No início dos anos 1900, outro fato importante foi o estabelecimento da natureza quântica da luz por Planck e Einstein, que permitiu a compreensão de que o Universo é permeado

¹ "[...]it is necessary to imagine that all space is crossed by the beams similar to beams of the X-ray, but considerably more penetrating."(CURIE, 1898 apud DORMAN; DORMAN, 2014, p. 5).

² "[...] If these beams exist, their source may be the Sun, and in that case will be various at midnight and at midday. However I could not find out it"(CURIE, 1899 apud DORMAN; DORMAN, 2014, p. 5).

³ "Experiments which will be made in the future, may be, will show that formation of ions in air deprived of any pollution, is caused by radiation which arises out of our atmosphere to similarly X-ray or cathodic rays, but possesses considerably bigger penetrating ability."(WILSON, 1901 apud DORMAN; DORMAN, 2014, p. 4).

por processos não-térmicos de energias extremas, que se estendem muito além da faixa óptica (SPURIO, 2018) .

A hipótese da origem terrestre dos raios cósmicos começou a ser questionada quando, em 1907, Artur Stewart Eve encontrou o mesmo nível de radiação no continente e no Oceano Atlântico, indicando que a crosta terrestre não deveria ser a fonte dessa radiação. Em 1909, Theodor Wulf observou que a ionização no topo da torre Eiffel diminuía muito menos do que o previsto pela teoria dominante na época. Simultaneamente, Domenico Pacini realizou experimentos entre 1910 e 1911 comparando medições na superfície e submersas no mar, nesses experimentos constatou que a ionização a 3 metros de profundidade era aproximadamente 20% menor que na superfície, concluindo que uma parte significativa da radiação provinha de fora da crosta terrestre.

A prova definitiva da origem extraterrestre, veio com os voos de balão de Victor Franz Hess em 1912 (Figura 3). Utilizando eletroscópios herméticos para garantir que a pressão dentro do instrumento fosse constante, Hess demonstrou que, após uma leve queda inicial até 1000 metros, a ionização aumentava drasticamente com a altitude, chegando a ser várias vezes maior a 5000 metros do que ao nível do mar. Hess concluiu que uma "radiação de ultra-gama" penetrava a atmosfera vinda de cima. Uma curiosidade trazida por Dorman e Dorman (2014) é que Franz Linke tinha obtido resultados semelhantes em vôos para seu doutorado entre 1900 e 1903, sua tese não foi publicada, mas relatórios indicam que, mesmo com os resultados corretos, ele chegou à conclusão errada influenciado pela queda inicial da ionização. Os resultados de Hess foram confirmados e ampliados pelas medições de Werner Kolhörster entre 1913 e 1914, que mediu um aumento de ionização ainda mais expressivo acima dos 9000 metros.

Figura 3 – Victor Hess em um de seus vôos, entre 1911 e 1912



Fonte: Breisky (2012).

A aceitação final da origem extraterrestre ocorreu apenas após os experimentos de Millikan e Cameron em lagos. Ao mergulhar instrumentos em diferentes profundidades eles provaram que a atmosfera atuava como um absorvedor equivalente à água, confirmando que a radiação vinha do espaço. A disseminação do uso do termo "Raios Cósmicos" para se referir à "radiação de altitude" de Hess é, muitas vezes, atribuída a Millikan⁴ em 1926⁵. A adoção do termo "raios" vem da hipótese sustentada por Millikan de que essa radiação primária era composta por fótons de altas energias (raios gama), resultantes do processo de nucleossíntese. Entre 1933 e 1934, Luis Alvarez e Arthur Compton identificaram que a maioria dos raios cósmicos medidos por eles tinham carga elétrica positiva, mas foi somente na década de 1950 que experimentos em altas altitudes mostraram a composição dos raios cósmicos: a maior parte dessa radiação era composta por prótons, seguido por $\sim 10\%$ de núcleos de He, $\sim 1\%$ de núcleos mais pesados, $\sim 10\%$ de elétrons e menos de 1% de raios gama.

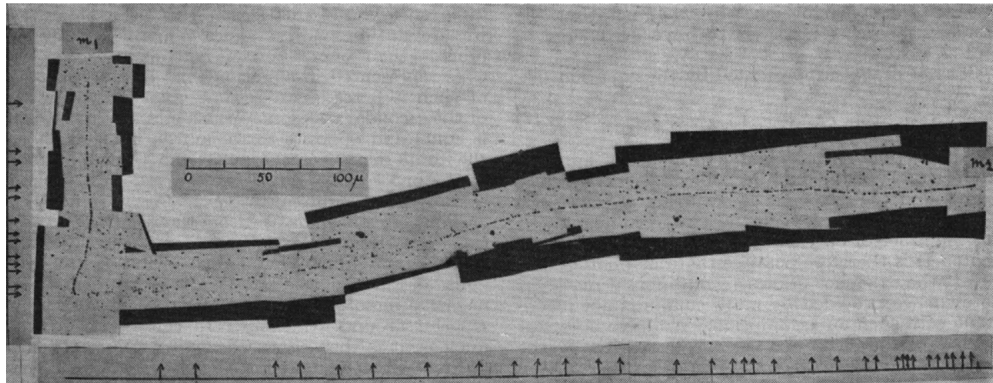
A investigação sobre a natureza particulada dos raios cósmicos culminou em uma das descobertas mais fundamentais da física do século XX: a detecção da antimatéria, quando Carl David Anderson detectou uma partícula de massa compatível com a do elétron, porém com trajetória oposta, indicando carga positiva. A esse "elétron positivo" foi dado o nome de pósitron. Essas descobertas inauguram a era da física de partículas de altas energias. Em 1936, a Academia Real das Ciências da Suécia laureou Victor Hess com o Prêmio Nobel de Física pela descoberta da radiação cósmica. Com o passar dos anos foram descobertas muitas outras partículas subatômicas nos raios cósmicos: múons, hiperons, mésons K, culminando na descoberta do pión (Méson π) (Figura 4), pelo brasileiro César Lattes utilizando a técnica de emulsões nucleares que rendeu o Nobel de Física de 1950 para seu grupo. A Física de partículas e a astrofísica andavam lado a lado até meados da década de 1950. Foi nessa década também que, após a 2ª Guerra Mundial, a janela de rádio foi aberta, expandindo ainda mais o horizonte observacional (SPURIO, 2018).

As décadas de 1960 e 1970 foram marcadas pela descoberta da radiação cósmica de fundo (*Cosmic Microwave Background* (CMB)), de pulsares, quasares, e pelo desenvolvimento de novos detectores e o uso de foguetes para superar a barreira da atmosfera terrestre. O avanço da tecnologia no decorrer do século XX permitiu o desenvolvimento de novas formas de obter informações sobre objetos astrofísicos. Transformando a astrofísica moderna em uma astronomia multimessageira, combinando a detecção de fótons, raios cósmicos, neutrinos e ondas gravitacionais para investigar os ambientes mais violentos do universo, conseguindo assim uma caracterização mais ampla do Universo que nos cerca (SPURIO, 2018). A Supernova SN 1987A, resultado do colapso de uma estrela na Grande Nuvem de Magalhães é tida como

⁴ Algumas fontes, como Chalmers (2020), afirmam que o termo "*kosmische strahlung*" foi utilizado por Albert Gockel, porém, em suas publicações ele manteve o termo "*durchdringende Strahlung*" que pode ser traduzido como radiação penetrante (GOCKEL, 1909; GOCKEL, 1910; GOCKEL, 1911).

⁵ These rays [...] do not occur from our atmosphere and consequently can be rightfully named by 'cosmic rays', this most descriptive and most suitable name" (MILLIKAN, 1926 apud DORMAN; DORMAN, 2014, p. 13)

Figura 4 – Chapas fotográficas da descoberta do Méson π



Fonte: Lattes *et al.* (1947).

marco inicial da astronomia multimessageira, antes que a luz visível da explosão atingisse a Terra, detectores subterrâneos (Kamiokande-II, IMB e Baksan) registraram uma rajada de neutrinos vindos daquela região (DORMAN; DORMAN, 2014).

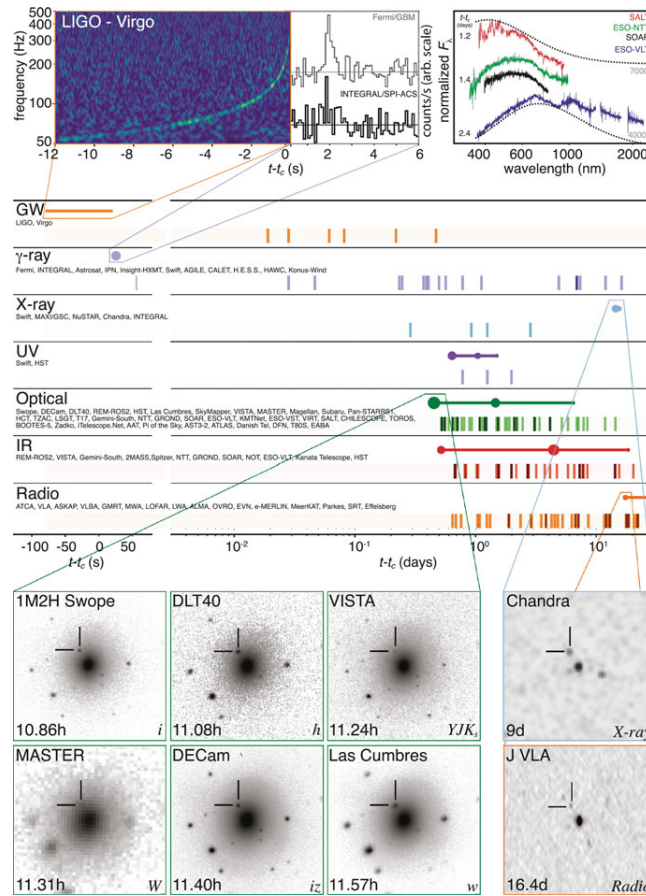
Para Spurio (2018), o evento de maior destaque da astronomia multimessageira aconteceu em 17 de agosto de 2017 (Figura 5), com a primeira detecção conjunta de ondas gravitacionais e radiação eletromagnética provenientes da mesma fonte astrofísica. Às 12:41:04 UTC, a rede de interferômetros LIGO e Virgo registrou o sinal GW170817, gerado pela fusão de duas estrelas de nêutrons, cerca de 2 s depois, os satélites Fermi e INTEGRAL detectaram uma explosão de raios gama (*Gamma-Ray Burst* (GRB)), designada GRB 170817A, vinda da mesma direção. Esse evento desencadeou a maior campanha observacional da história, mobilizando cerca de 70 observatórios terrestres e espaciais que identificaram a contrapartida óptica e infravermelha, denominada kilonova (AT 2017gfo), na galáxia NGC 4993, a 40 Mpc de distância. A espectroscopia posterior revelou a assinatura de elementos pesados recém-sintetizados, como ouro e platina.

2.2 Raios Gama

Raios Cósmicos (RCS) são partículas que vêm de fora da Terra, como um dos objetivos centrais deste trabalho é a separação entre chuviros atmosféricos iniciados por raios gama daqueles iniciados por hádrons, adotaremos o termo "raios cósmicos" ao tratar de partículas carregadas, geralmente núcleos atômicos, que, devido à interação com campos magnéticos, viajam em trajetórias difusas e imprevisíveis, perdendo informação direcional. O termo "raios gama" se refere aos fótons de mais altas energias (acima de 0,511 MeV) que viajam em linha reta a partir da fonte, sem sofrer deflexão por campos magnéticos, preservando a informação direcional e permitindo o imageamento direto dos aceleradores cósmicos no qual foram gerados, podemos ter uma visão geral ao olharmos para a figura 6.

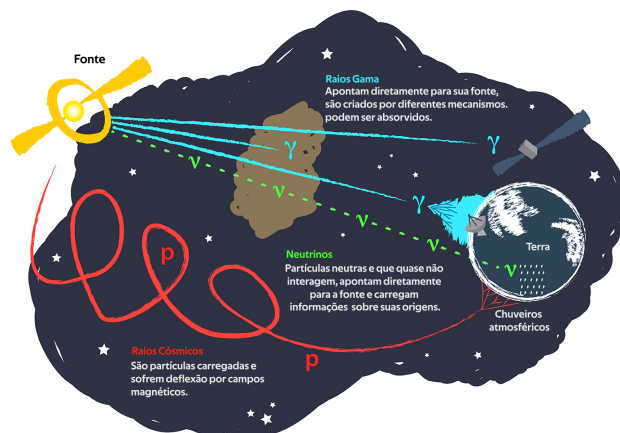
A primeira detecção de raios gama extraterrestres ocorreu apenas em 1965 com o satélite Explorer XI, seguido por avanços significativos com as missões SAS-2 (1972) e COS-B

Figura 5 – Fusão de duas estrelas de nêutrons observada em diferentes sinais, considerado um dos eventos de maior destaque da astronomia multimessageira



Fonte: Spurio (2018).

Figura 6 – Trajetória e interação das partículas no espaço



Fonte: Adaptado de Filipović, Payne e Tothill (2021).

(1975), que mapearam o plano galáctico e identificaram as primeiras fontes discretas de raios gama no céu, o primeiro telescópio óptico construído com o propósito de ser usado para a astronomia de raios Gama foi o Whipple, um telescópio de 10 m de diâmetro instalado no Monte Hopkins, em 1968, o desenvolvimento da técnica de imageamento Cherenkov permitiu a de-

tecção da Nebulosa do Caranguejo (Weekes *et al.*, 1989), hoje usada como vela padrão na astronomia de raios gama.

A astronomia de raios gama é tida como a "última janela do espectro eletromagnético a ser aberta" (AHARONIAN, 2004, p 1). Uma janela enorme, que cobre pelo menos 14 décadas na energia, indo de $E \approx 0.511 \times 10^6$ eV até $\gtrsim 10^{20}$ eV. Sendo o limite inferior caracterizado pela energia da aniquilação de um par elétron-pósitron e o limite superior representando as energias das partículas mais altas já observadas. Dentro desses limites, Aharonian (2004) divide a astronomia de raios gama em 5 áreas:

Baixa Energia (*Low Energy* (LE)): até 30 MeV.

Alta Energia (*High Energy* (HE)): 30 MeV a 30 GeV.

Muito Alta Energia (*Very-High Energy* (VHE)): 30 GeV a 30 TeV.

Ultra Alta Energia (*Ultra-High Energy* (UHE)): 30 TeV a 30 PeV.

Extremamente Alta Energia (*Extremely-High Energy* (EHE)): acima de 30 PeV.

2.2.1 Produção e Propagação

Uma das formas de emitir fótons na natureza é através do aquecimento da matéria, olhando para o espectro de emissão de um corpo negro em função da energia, dado pela distribuição de Planck, temos:

$$I_{BB} = \frac{2E_{ph}^3}{h^3 c^2} \left[\frac{1}{\exp(E_{ph}/kT) - 1} \right] [ergs^{-1} cm^{-2} str^{-1}]. \quad (1)$$

Onde:

I_{BB} : Intensidade específica do corpo negro.

E_{ph} : Energia do fóton ($h\nu$).

h : Constante de Planck (6.625×10^{-27} erg s⁻¹)

c : Velocidade da luz.

k : Constante de Boltzmann (1.38×10^{-16} erg K⁻¹).

T : Temperatura absoluta (K).

A energia do fóton é dada pela equação:

$$\langle E_{ph} \rangle \approx 2,7kT \approx 2,3 \times 10^{-10} \left(\frac{T}{K} \right) [MeV]. \quad (2)$$

Para que um fóton de 1 GeV seja criado, seria necessária uma temperatura da ordem de 10^{13} K, uma fonte aquecida a essa temperatura teria uma densidade de fótons de $n_{ph} \approx 5 \times 10^{46}$ cm⁻³, levando a um caminho livre médio de:

$$\lambda_{ph} \approx (n_{ph} \sigma_{\gamma\gamma})^{-1} \approx 2 \times 10^{-23} [cm], \quad (3)$$

Sendo $\sigma_{\gamma\gamma} \approx 10^{-24} \text{ cm}^2$ a aproximação da seção de choque de Thomson.

Ou seja, a própria fonte torna-se opaca, toda a energia vai para a produção de pares ($\gamma + \gamma \rightarrow e^+ + e^-$). A geração de matéria faz com que haja um aumento da pressão na fonte, que se expande e, eventualmente, torna-se transparente novamente.

Portanto, Aharonian (2004) pontua que a detecção de partículas de altas energias indica a existência de mecanismos astrofísicos de natureza não térmica e de alta eficiência. A geração de radiação gama em VHE está diretamente associada a partículas carregadas, aceleradas a velocidades relativísticas.

2.2.1.1 Mecanismos de Aceleração

Embora existam modelos *Top-Down*, que propõem que partículas de altas energias podem surgir a partir de defeitos topológicos, cordas cósmicas ou outros mecanismos exóticos, Spurio (2018) considera que esses modelos criam mais problemas do que resolvem, visto que até hoje não foram encontradas evidências da existência desses objetos. Nesse trabalho abordaremos os modelos *Bottom-Up*, onde partículas carregadas estáveis são aceleradas progressivamente por meio de processos, principalmente, eletromagnéticos em ambientes extremos do Universo, focando nos processos capazes de produzir VHE na faixa de operação do LST, entre 20 e 150 GeV.

O modelo *Bottom-Up* satisfaz as condições restritivas geométricas e radiativas citadas por Ptitsyna e Troitsky (2010):

Geometria: a partícula em aceleração devido à ação de um campo magnético deve continuar nos domínios da fonte, como postulado pelas condições de Hillas e ilustrado na figura 7:

$$R_o \gtrsim R_L \Rightarrow R_o \gtrsim \frac{1}{Z} \left(\frac{B_o}{1 \text{ nT}} \right)^{-1} \left(\frac{E}{10[EeV]} \right) [kpc]. \quad (4)$$

Onde:

R_o : Tamanho característico da região de aceleração do objeto astronômico.

R_L : Raio de Larmor da partícula carregada.

Z : Número atômico da partícula acelerada.

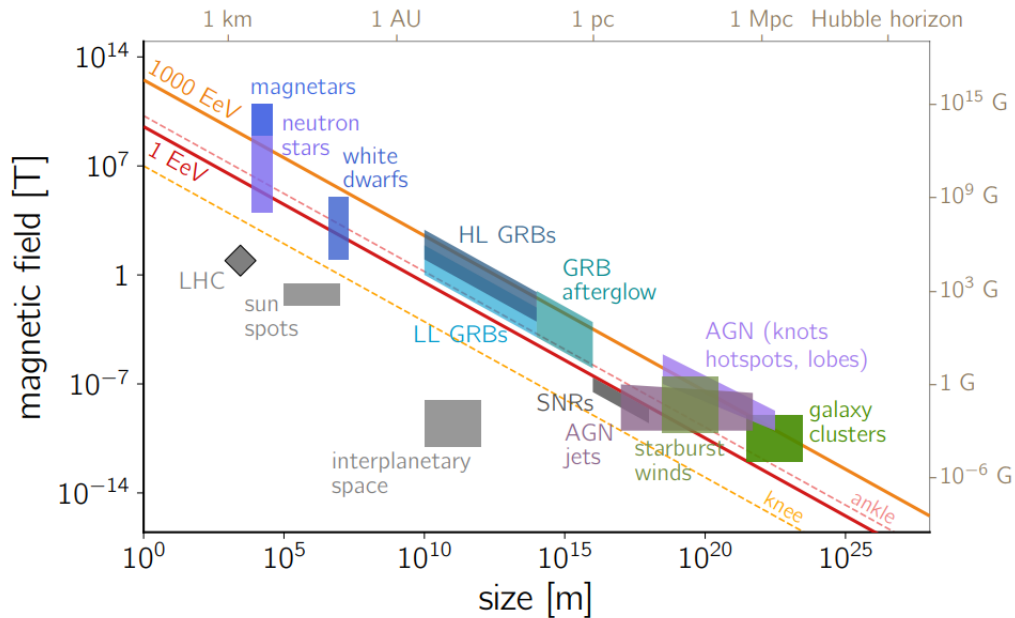
B_o : Intensidade do campo magnético na região.

E : A energia final da partícula.

Potência: a fonte deve possuir energia suficiente para transferir a energia necessária às partículas aceleradas. Se assumirmos que a energia é fornecida por um campo magnético, chegamos ao Limite de Hillas-Lovelace:

$$P_{B,\min} \gtrsim u_B R_o^2 v \Rightarrow P_{B,\min} \gtrsim \frac{10^{37} \eta^2}{\beta Z^2} u_B \left(\frac{E_{\max}}{10[EeV]} \right)^2 [Js^{-1}]. \quad (5)$$

Figura 7 – Critério de Hillas, ilustrando diferentes classes de aceleradores: O eixo vertical representa a intensidade do campo magnético enquanto o eixo horizontal representa o tamanho do objeto, as linhas sólidas representam o limite de energia na qual a partícula "escapa" dos limites do objeto, enquanto as linhas tracejadas indicam as regiões do joelho e tornozelo no espectro.



Fonte: Batista (2025).

Onde:

$P_{B,\min}$: Potência mínima exigida da fonte.

u_B : Densidade de energia magnética ($B^2/8\pi$).

R_0 : Tamanho característico ou raio da região de aceleração.

v e β : Velocidade da expansão/jato e a velocidade normalizada (v/c).

η : Parâmetro de eficiência.

Z : Número atômico da partícula acelerada.

E_{\max} : Energia máxima final da partícula.

Ganho vs perda de energia: A taxa de energia perdida pela partícula, seja por radiação ou interação com o meio não deve ser maior que a energia ganha pela partícula.

Emissividade: a quantidade e potência das fontes deve explicar o fluxo de partículas de altas energias observados.

Composição: O fluxo de partículas secundárias (fótons, neutrinos e raios cósmicos de baixa energia) gerado pelo processo de aceleração não deve exceder os limites de fluxo já observados experimentalmente.

Em resumo, para que ocorra a emissão de raios gama de altas energias, é pré-requisito a existência de populações de partículas carregadas sendo aceleradas. Essa aceleração pode acontecer de duas formas:

Direta (ou indutiva): A aceleração da partícula acontece, como o próprio nome indica, de forma direta pela ação de campos elétricos ou magnéticos intensos. Existem estudos que

indicam que essa aceleração também pode acontecer quando a partícula é exposta à campos gravitacionais extremos, como ao redor de pulsares ou buracos negros (PEREIRA *et al.*, 2024).

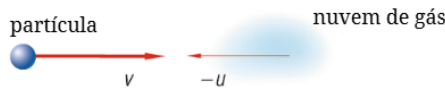
Difusa (ou estocástica): A aceleração ocorre por meio da interação estocástica da partícula com irregularidades do campo magnético, ondas de plasma, turbulência ou choques, essas interações sucessivas geram um ganho gradual de energia. Aqui destacam-se os mecanismos de Fermi de 1ª e 2ª ordem, propostos pelo físico Enrico Fermi em 1949 para explicar como partículas poderiam ganhar energia ao interagir com nuvens em regiões interestelares.

2.2.1.1.1 Mecanismo de Fermi de Segunda Ordem

Originalmente, Fermi propôs que as partículas ganham energia através de colisões com nuvens magnéticas ou de gás que se movem aleatoriamente no meio interestelar. A figura 8 mostra uma partícula de massa m e velocidade v colidindo com uma nuvem de massa $M \gg m$ e velocidade u .

Figura 8 – Mecanismo de Fermi de segunda ordem

caso 1:



caso 2:



Fonte: Grupen (2020).

Considerando que v e u são antiparalelas, ou seja, viajam em direções opostas, temos que:

$$\Delta E_1 = \frac{1}{2}m(v + u)^2 - \frac{1}{2}mv^2 = \frac{1}{2}m(2uv + u^2). \quad (6)$$

A partícula ganha energia, já quando v e u são paralelas, com ambas viajando na mesma direção, a partícula perde energia:

$$\Delta E_2 = \frac{1}{2}m(v - u)^2 - \frac{1}{2}mv^2 = \frac{1}{2}m(-2uv + u^2). \quad (7)$$

Na média, o ganho de energia da partícula é dado por:

$$\Delta E = \Delta E_1 + \Delta E_2 = mu^2, \quad (8)$$

o que leva a um ganho relativo de energia:

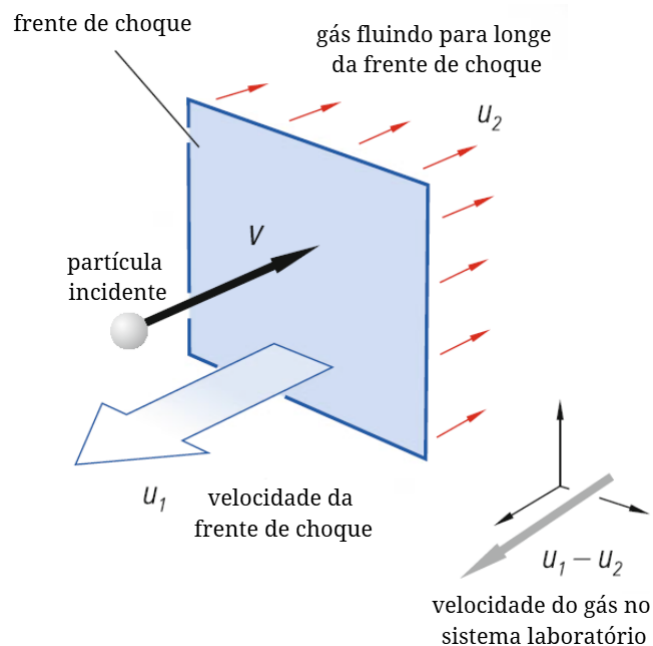
$$\frac{\Delta E}{E} = 2 \frac{u^2}{v^2}. \quad (9)$$

O termo quadrático na velocidade dá o nome ao mecanismo, que não é muito eficiente, visto que a velocidade da nuvem costuma ser relativamente baixa ($\frac{u}{c} \approx 10^{-4}$), é necessária uma escala temporal bastante grande para acelerar partículas até altas energias. Além disso, para compensar as perdas, a partícula precisa ser injetada na nuvem com uma energia mínima, essa energia pode ser fornecida pelo Mecanismo de Fermi de Primeira Ordem.

2.2.1.1.2 Mecanismo de Fermi de Primeira Ordem

Também chamado de aceleração por ondas de choque, ocorre principalmente em supernovas onde a camada ejetada durante a explosão age como uma onda de choque no meio interestelar. Segundo Longair (2011), esse é o paradigma mais aceito desde a década de 1970, pois resulta em uma lei de potência com índice espectral bem próximo ao observado.

Figura 9 – Mecanismo de Fermi de primeira ordem



Fonte: Grupen (2020).

Consideremos a Figura 9: uma partícula viajando em uma região onde encontra uma onda de choque que se move-se com velocidade supersônica u_1 , enquanto o gás atrás do choque se desloca a uma velocidade subsônica u_2 . Ao colidir com essa frente e ser refletida, a partícula de velocidade v experimenta um ganho de energia de:

$$\Delta E = \frac{1}{2}m(v + (u_1 - u_2))^2 - \frac{1}{2}mv^2 = \frac{1}{2}m(2v(u_1 - u_2) + (u_1 - u_2)^2), \quad (10)$$

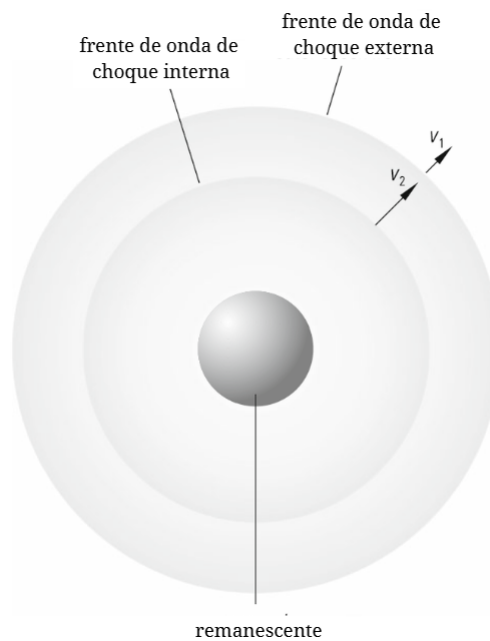
como o termo linear ($v \gg u_1, u_2, u_1 > u_2$) domina, o ganho relativo de energia é dado por:

$$\frac{\Delta E}{E} \approx \frac{2(u_1 - u_2)}{v}. \quad (11)$$

Ao tratarmos desse sistema de forma relativística, inserindo o momento e o ângulo de espalhamento da partícula durante a interação e aplicando os fatores de Lorentz, chegamos à um ganho linear de energia de:

$$\frac{\Delta E}{E} = \frac{4}{3} \frac{u_1 - u_2}{c}. \quad (12)$$

Figura 10 – Aceleração por múltiplas reflexões entre ondas de choque



Fonte: Grupen (2020).

Na figura 10 ilustra-se um cenário que ocorre em uma supernova, no qual a partícula está presa entre duas ondas de choque, uma onda de choque mais interna com uma velocidade (v_2) muito maior que a velocidade da onda de choque mais externa (v_1), a partícula ao ser refletida pela onda de choque mais interna ganha energia:

$$\Delta E_1 = \frac{1}{2}m(v + v_2)^2 - \frac{1}{2}mv^2 = \frac{1}{2}m(v_2^2 + 2vv_2). \quad (13)$$

E a energia perdida pela reflexão na onda de choque mais externa é dada por:

$$\Delta E_2 = \frac{1}{2}m(v - v_1)^2 - \frac{1}{2}mv^2 = \frac{1}{2}m(v_1^2 - 2vv_1). \quad (14)$$

Em média, a partícula ganha uma energia

$$\Delta E = \frac{1}{2}m(v_1^2 + v_2^2 + 2v(v_2 - v_1)). \quad (15)$$

Os termos quadráticos podem ser desprezados, levando a:

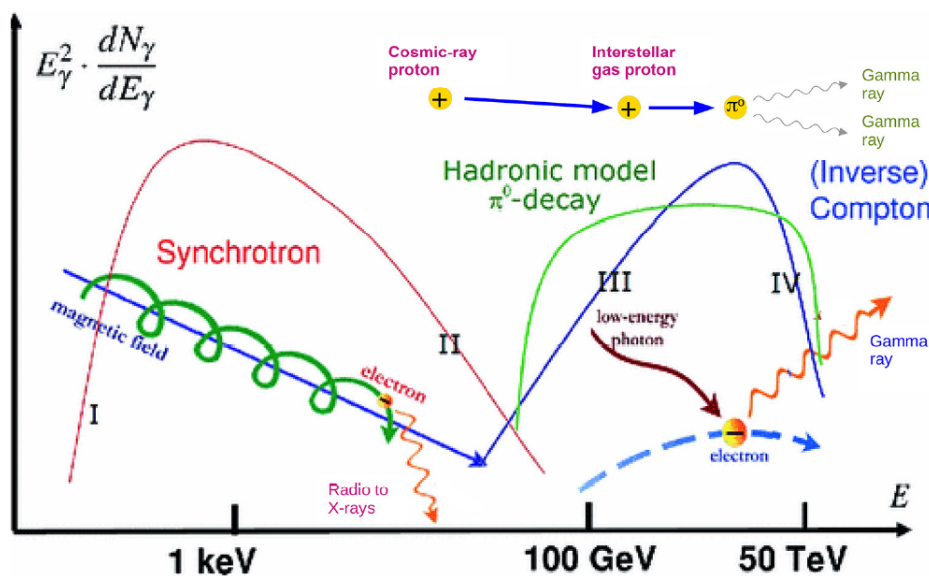
$$\Delta E \approx mv\Delta v, \quad \frac{\Delta E}{E} \approx 2\frac{\Delta V}{v}. \quad (16)$$

Resultando em um processo mais eficiente, podendo explicar acelerações de partículas até a ordem de 100 TeV (GRUPEN, 2020) em um intervalo de tempo significativamente menor do que o Mecanismo de Fermi de Segunda Ordem (SPURIO, 2018).

2.2.1.2 Geração de Raios Gama de Altas Energias

Uma vez aceleradas, essas partículas interagem com a matéria ou campos de radiação circundantes, convertendo parte de sua energia cinética em radiação gama através de processos não-térmicos. Como vemos na figura 11, a aceleração síncrotron domina a aceleração de partículas com energias mais baixas, enquanto três processos dominam a emissão na faixa de energia acima de 20 GeV: Bremsstrahlung e Espalhamento Compton Inverso como processos leptônicos e decaimento de mésons neutros como processo hadrônico.

Figura 11 – Espectro de emissão Gama: Bremsstrahlung não é representado por ser subdominante no intervalo



Fonte: Filipović, Payne e Tothill (2021).

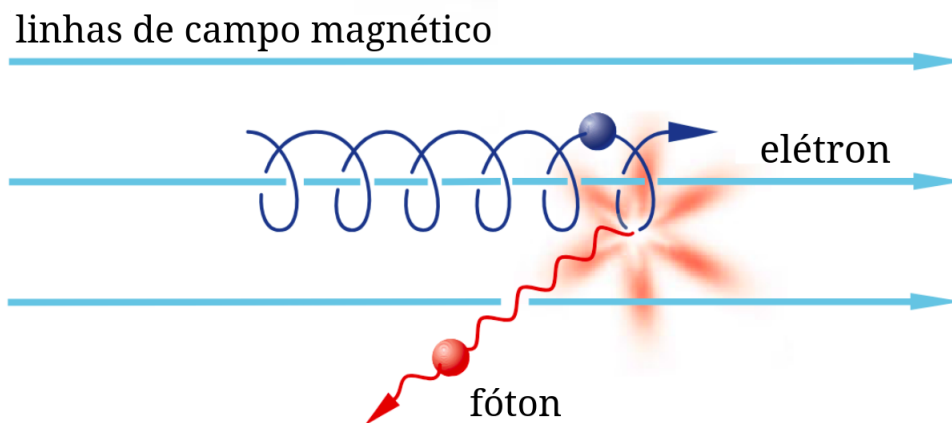
2.2.1.2.1 Síncrotron

A radiação síncrotron é uma forma de radiação emitida por partículas carregadas viajando em uma trajetória curvilínea ao redor de linhas de campos magnéticos homogêneos. Qualquer partícula carregada pode gerar radiação síncrotron, mas para a radiação síncrotron produzida no contexto cósmico a predominância é de radiação emitida por elétrons. Apresenta um espectro contínuo e a potência irradiada P é dada por:

$$P \sim E^2 B^2, \quad (17)$$

sendo E a energia do elétron e B a intensidade do campo magnético (GRUPEN, 2020).

Figura 12 – Produção de radiação síncrotron pela deflexão de partículas carregadas em um campo magnético



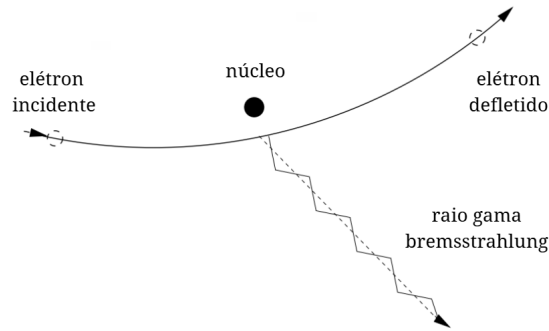
Fonte: Grupen (2020).

O espectro de emissão gama por radiação síncrotron é mais significativo para energias mais baixas, como visto na Figura 11, não sendo capaz de gerar diretamente fótons acima da ordem de GeV. Porém, como apontado por Longair (2011), essas energias extremas podem ser alcançadas através do mecanismo *Synchrotron Self-Compton*. Neste processo, a mesma população de elétrons ultra-relativísticos interage com os fótons síncrotron recém-gerados, transferindo momento e energia por meio do espalhamento Compton Inverso.

2.2.1.2.2 Bremsstrahlung

O mecanismo de *Bremsstrahlung*, ilustrado na figura 13, ocorre quando uma partícula carregada sofre uma aceleração ao ser defletida pelo campo elétrico de outra carga, tipicamente um núcleo atômico, o termo "*Bremsstrahlung*" pode ser traduzido como "radiação de frenagem" (HORVATH, 2022). E assume que haverá radiação eletromagnética sempre que uma carga está em movimento.

Figura 13 – Emissão de raio gama por bremsstrahlung



Fonte: Weekes (2003).

Uma das partículas mais abundantes e de fácil aceleração são os elétrons. Um elétron é defletido de sua trajetória quando encontra um campo eletromagnético intenso, emitindo um fóton. Considerando um parâmetro de impacto b (a distância entre o centro da fonte para não ser defletido). Aplicando a Lei de Newton, a desaceleração da partícula devido ao campo coulombiano será:

$$\mathbf{F} = m\mathbf{a} \approx -\frac{Ze^2}{x^2}. \quad (18)$$

Ou seja, a magnitude da aceleração é $a \approx Ze^2/mx^2$, a desaceleração do elétron é mais significativa quando ele está uma distância x mais próxima do núcleo, podemos então considerar como um pulso de curta duração. A potência emitida durante esse tempo ($\Delta T \approx 2b/v$) pode ser calculada usando a Fórmula de Larmor:

$$P = -\frac{dE}{dt} = \frac{2e^2}{3c^3}a^2. \quad (19)$$

Inserindo a aceleração calculada no caso de *bremsstrahlung*:

$$P = -\frac{dE}{dt} = \frac{2e^2}{3c^3} \left(\frac{Ze^2}{mx^2} \right)^2. \quad (20)$$

A energia irradiada num único encontro pode ser estimada como o produto da potência pelo tempo de interação:

$$P\Delta t = \frac{2e^2}{3c^3} \left(\frac{Ze^2}{mb^2} \right)^2 \times \frac{2b}{v} = \frac{4}{3} \frac{Z^2e^6}{3c^3m^2b^3v}. \quad (21)$$

E a energia máxima irradiada pelo elétron em um intervalo de frequência é:

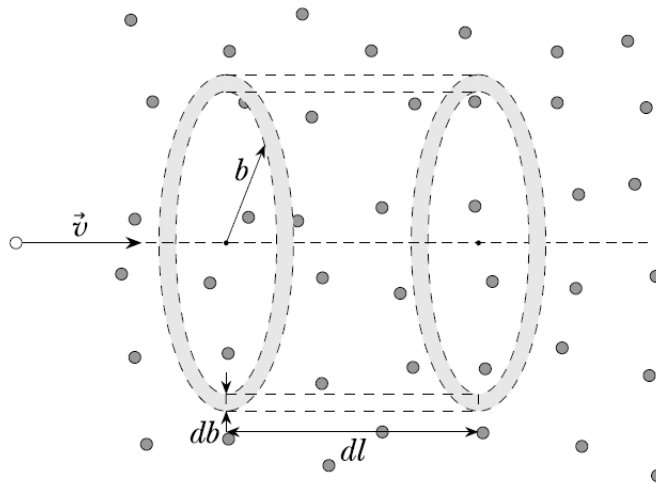
$$\frac{P\Delta t}{\Delta\nu} \approx \frac{P\Delta t}{\nu_{\max}} = \frac{16}{3} \frac{Z^2e^6}{c^3m^2b^2v^2}. \quad (22)$$

Na prática não há a colisão de apenas um elétron, mas de uma nuvem de elétrons e íons (Figura 14), para efeitos de simplificação, vamos supor que todos esses íons tem carga Z .

Entre b e $b + db$ existe uma quantidade de íons igual a $2\pi n_Z v b db$, sendo n_Z a densidade de íons. A emissividade total dessa nuvem é dada integrando as interações ao longo do raio da nuvem:

$$\begin{aligned}
 I &= 2\pi n_e n_Z v \int_{b_{\min}}^{b_{\max}} \frac{16}{3} \frac{Z^2 e^6}{c^3 m^2 b^2 v^2} b db \\
 &= \frac{16}{3} \frac{\pi Z^2 e^6 n_e n_Z}{m^2 c^3 v} \int_{b_{\min}}^{b_{\max}} \frac{db}{b}. \\
 ZAS &= \frac{16\pi Z^2 e^6 n_e n_Z}{3m^2 c^3 v} \ln \left(\frac{b_{\max}}{b_{\min}} \right)
 \end{aligned} \tag{23}$$

Figura 14 – Geração de raios gama via bremsstrahlung pela colisão entre nuvens



Fonte: Horvath (2022).

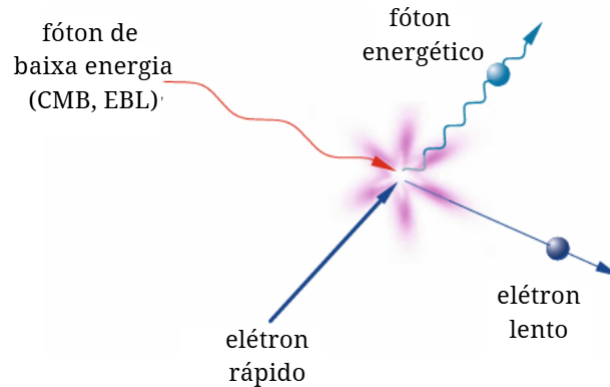
A emissão espectral devido ao *bremsstrahlung* apresenta um fluxo de fótons que obedece à mesma lei de potência dos elétrons incidentes. Assim, se a população de elétrons apresenta um espectro $dN_e/dE \propto E^{-\alpha}$, os raios gama gerados terão um espectro $dN_\gamma/dE \propto E^{-\alpha}$ (AHARONIAN, 2004).

2.2.1.2.3 Espalhamento Compton Inverso

Esse mecanismo é chamado de inverso pois, em vez dos fótons perderem energia, como acontece no Efeito Compton tradicional, quem transfere energia cinética são elétrons relativísticos, acelerando fótons de fundo até a faixa de raios gama (Figura 15).

Horvath (2022) aponta que a dedução matemática é análoga à do Efeito Compton, sendo necessária uma Transformada de Lorentz para o referencial do elétron. Como descrito por (LONGAIR, 2011), se considerarmos um elétron ultra-relativístico com fator de Lorentz $\gamma = E_e mc^2 \gg 1$ colidindo com um fóton de baixa energia no referencial do laboratório (S). O

Figura 15 – Ilustração do mecanismo de emissão de raios gama via espalhamento Compton inverso



Fonte: Grupen (2020).

centro de momento do sistema de referência será bem próximo ao do elétron, se o fóton tiver uma energia $\hbar\omega$ e o ângulo de incidência θ em S , a energia no referencial S' , será:

$$\hbar\nu = \gamma\hbar\omega[1 + (v/c)\cos\theta]. \quad (24)$$

O ângulo de incidência θ' no referencial S' pode ser relacionado ao θ em S através da fórmula:

$$\sin\theta' = \frac{\sin\theta}{\gamma[1 + (v/c)\cos\theta]}; \cos\theta' = \frac{\cos\theta + v/c}{[1 + (v/c)\cos\theta]}. \quad (25)$$

Como $\hbar\omega' \ll m_e c^2$, a interação no referencial de repouso do elétron (S') ocorre no Regime de Thomson, onde o espalhamento é essencialmente elástico e a seção de choque de interação é a seção de choque de Thomson (σ_T), que é independente da energia. A taxa de perda de energia do elétron em S' pode ser expressa como:

$$-\left(\frac{dE}{dt}\right)' = \sigma_T c u'_{rad}. \quad (26)$$

u'_{rad} representa a densidade de energia da radiação no referencial do elétron (S'). Ao transformar essa taxa de volta para o referencial do laboratório (S), e considerando que os fótons espalhados em S' possuem uma distribuição aproximadamente isotrópica, a energia final média do fóton ($\hbar\omega_s$) sofre um novo fator de aumento proporcional a γ .

A energia final do fóton no referencial do laboratório S ($\hbar\omega_f$), deve ser obtida através de uma transformação inversa de Lorentz. Ao retornar para S , o fóton ganha um segundo fator de deslocamento Doppler e a máxima energia transferida é de:

$$(\hbar\omega)_{max} \approx 4\gamma^2 \hbar\omega_0, \quad (27)$$

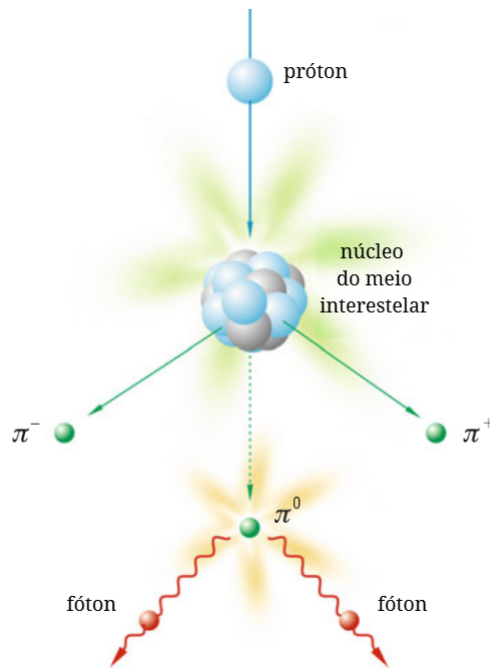
que, ao ser integrada sobre todos os ângulos de incidência e espalhamento, resulta em um ganho médio de energia dos fótons proporcional ao quadrado do fator de Lorentz:

$$\hbar\bar{\omega} \approx \frac{4}{3}\gamma^2\hbar\omega_0. \quad (28)$$

O fator γ^2 explica a eficiência do mecanismo, sabe-se que algumas fontes astrofísicas, como AGNs ou SNRs, são capazes de acelerar elétrons com $\gamma \sim 100 - 1000$, tais elétrons podem acelerar fótons diretamente para a faixa de TeV (SPURIO, 2018).

2.2.1.2.4 Decaimento de mésons neutros

Figura 16 – Produção de píon neutro por interação próton-núcleo e posterior decaimento em um par fóton-fóton



Fonte: Grupen (2020).

Diferente dos elétrons, os prótons, por terem grande massa, não perdem energia de forma significativa por radiação direta. Nesses casos, a produção de raios gama ocorre, principalmente, através de colisões inelásticas de prótons com a matéria do meio interestelar (Figura 16), podendo ser outros prótons ou núcleos mais pesados, gerando partículas secundárias (SPURIO, 2018).

$$p + nucleo \rightarrow \pi^{\pm}, \pi^0, K^{\pm}, K^0, p', nucleo', n, \dots \quad (29)$$

Outra forma de produzir essas partículas é a partir da interação de prótons ultra-relativísticos com fótons de baixa energia presentes no meio (γ_e), através da ressonância Δ^+ :

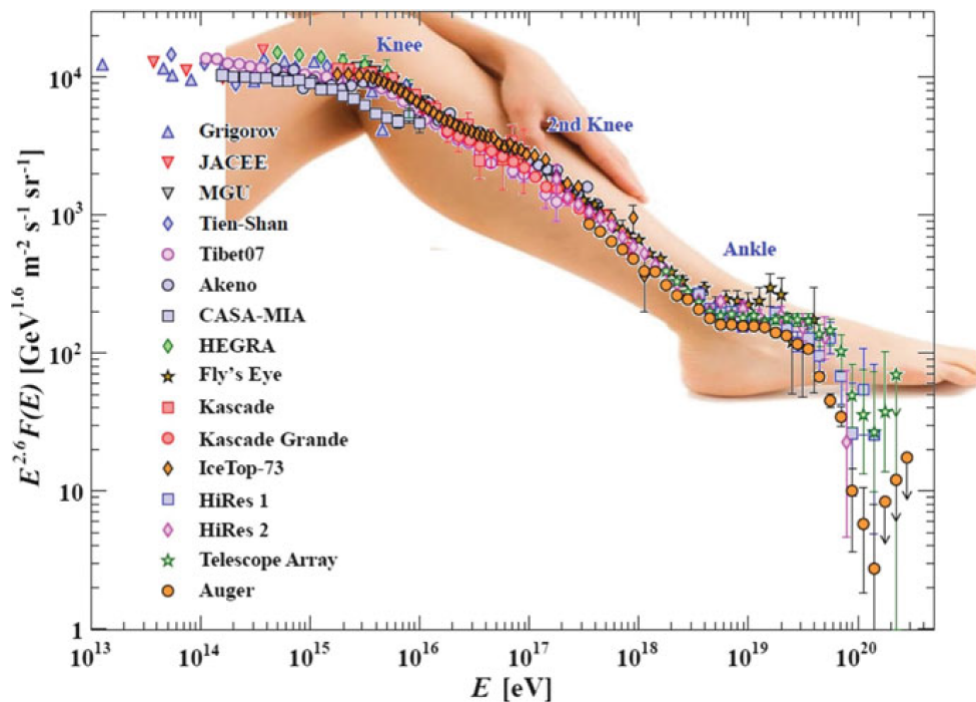


O méson de maior interesse aqui é o π^0 , visto que decai quase que instantaneamente ($8,4 \times 10^{-17} \text{ s}$) em dois fótons (em $\sim 98\%$ dos casos (ZYLA P A *et al.*, 2020)). Outra informação importante, colocada por Aharonian (2004), é que para energias muito acima de 1 GeV , o índice espectral dos fótons tende a ser o mesmo que o dos prótons que os geraram.

2.2.2 Fontes Astrofísicas

É consenso que a origem exata dos raios cósmicos ainda é um dos problemas sem solução na astrofísica (AHARONIAN, 2004; GRUPEN, 2020; SPURIO, 2018). A figura 17 tipicamente apresenta o espectro de raios cósmicos que chegam até a Terra, nela é possível perceber que existem duas estruturas em evidência: a região do joelho (*knee*) e a região do tornozelo (*ankle*). É bem provável que as partículas abaixo do joelho tenham origem galáctica, enquanto partículas com energias extremas, acima do tornozelo, tenham origem extragaláctica (AAB *et al.*, 2017).

Figura 17 – Fluxo observado de raios cósmicos de alta energia: eventos de ultra-alta energia são raros

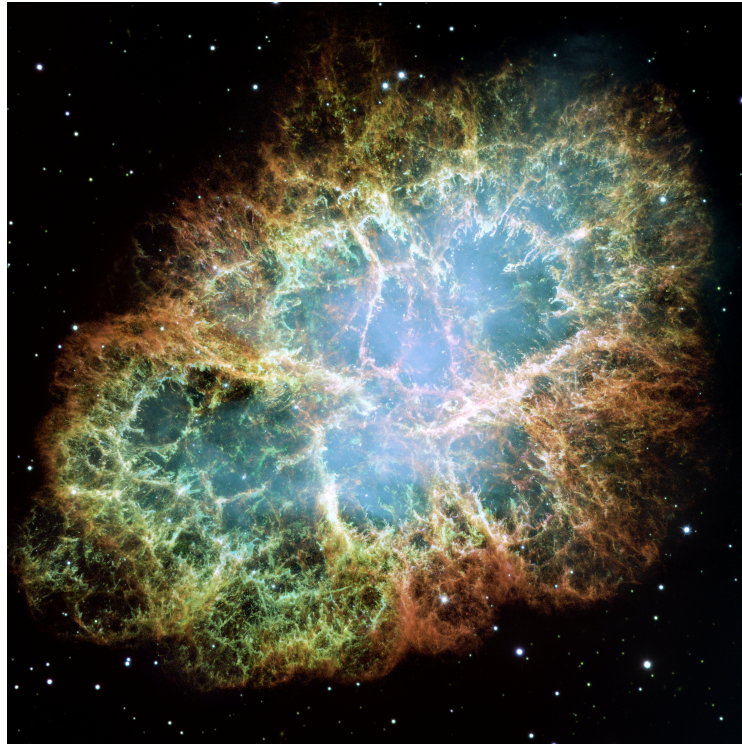


Fonte: Spurio (2018).

Diversas classes de objetos atendem aos critérios de Hillas (Figura 7) e possuem as condições extremas necessárias para atuar como aceleradores de partículas, algumas das que mais nos interessam neste trabalho são:

1. Remanescentes de Supernovas (SNRs): São considerados os principais aceleradores de raios cósmicos galácticos até a região do joelho ($\sim 10^{15} eV$). Uma Supernova acontece quando o processo de fusão de uma estrela chega até o Ferro, nesse momento a pressão gerada pela fusão não é suficiente para vencer a gravidade e a estrela implode, ejetando parte de sua massa para o espaço (Figura 18). Durante esse processo também acontece a síntese de elementos pesados. As camadas ejetadas representam uma onda de choque com relação ao meio interestelar, podendo viajar a mais de $20000 km/s$, e acelerando partículas via mecanismo de Fermi, como visto na seção 2.2.1.1.2 (SPURIO, 2018).

Figura 18 – Estrutura típica de um remanescente de supernova - Nebulosa do Caranguejo



Fonte: Spurio (2018).

2. Nebulosas de Vento de Pulsar (PWNe): Pulsares são estrelas de nêutrons superdensas ($\sim 6 \times 10^3 gcm^3$) que ficam após uma supernova (Figura 19), durante o processo de colapso da estrela o momento angular é conservado e o Pulsar apresenta períodos de rotação extremamente baixos, da ordem de milissegundos. O colapso da estrela também faz com que o campo magnético seja amplificado, alcançando $10^{12} G$ ($10^{15} G$ em magnetares), emitindo partículas via radiação síncrotron e Compton Inverso, enviando um "vento" de léptons altamente energético que ioniza o gás interestelar (SPURIO, 2018) (WEEKES, 2003).

Figura 19 – Pulsar na região central da Nebulosa do Caranguejo



Fonte: Spurio (2018).

3. Sistemas binários contendo um objeto compacto: Sistemas onde um objeto como buraco negro ou estrela de nêutrons tem uma estrela companheira. Nesses sistemas há um disco de acreção ao redor do objeto compacto, um plasma em rotação que gera campos eletromagnéticos extremos capazes de acelerar partículas a energias na ordem de $10^{19} eV$.

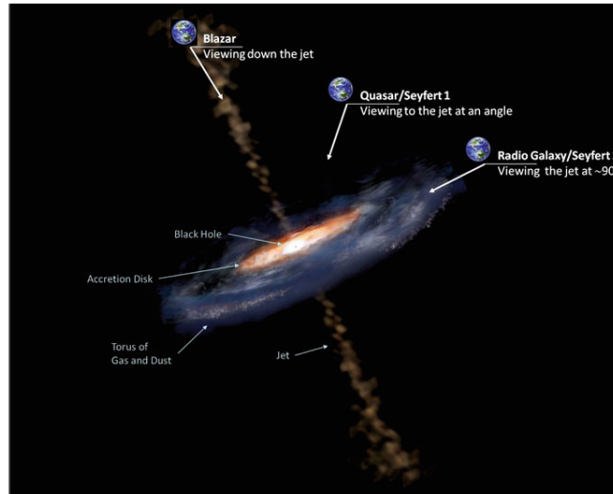
Figura 20 – Formação do disco de acreção em um sistema binário com objeto compacto



Fonte: Grupen (2020).

4. Núcleos Ativos de Galáxias (AGNs): Buracos negros supermassivos ativos no centro de galáxias podem ejetar jatos relativísticos capazes de acelerar partículas a energias extremas, além da aceleração devido aos fortes campos eletromagnéticos do disco de acreção. São os principais candidatos a fontes de UHECRs (LONGAIR, 2011).

Figura 21 – Núcleo Ativo de Galáxia (AGN), um buraco negro central é responsável pela alta luminosidade desses objetos

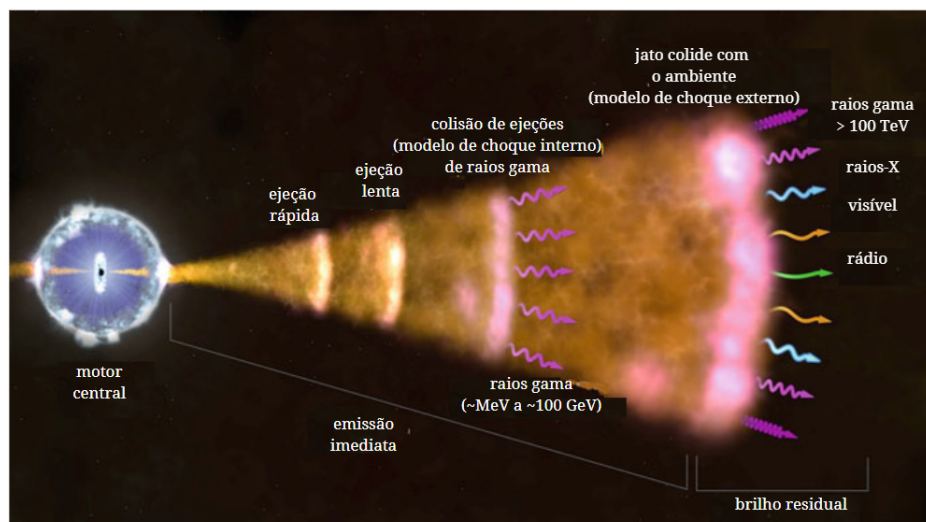


Fonte: Spurio (2018).

5. Explosões de Raios Gama (GRBs): São os eventos mais extremos conhecidos no Universo, liberam em questão de segundos mais energia do que o Sol produzirá em toda sua vida (~ 10 bilhões de anos), são os fenômenos mais brilhantes desde o *Big-Bang* (WEEKES, 2003). Explosões de Raios Gama são divididas em duas classes de acordo com a sua duração, os de longa duração ($\tau_{90} > 2$ s), são associados ao colapso de núcleos de estrelas massivas ($10M_{\odot}$) e tem uma duração média de ~ 30 s, enquanto os de curta duração ($\tau_{90} \leq 2$ s), são associados a fusão de objetos compactos, como binárias de estrelas de nêutrons e buracos negros, tendo duração média de 0,3 s.

Em ambos os casos a emissão de altas energias é descrita pelo "Modelo de Bola de Fogo" (*Fireball Model*), mostrado na figura 22, no qual um motor central ejeta jatos ultra-relativísticos, com fator de Lorentz $\sim 3 \times 10^2$.

Figura 22 – Modelo de bola de fogo para explosões de raios gama



Fonte: Spurio (2018).

2.2.2.1 Propagação

O espaço intergaláctico não é vazio, ele é permeado por campos de radiação difusa, e durante a propagação em distâncias cosmológicas, os raios gama podem interagir com esses fótons de fundo, resultando em uma atenuação do fluxo observado na Terra. O principal mecanismo atuante na faixa até 30 TeV é a produção de pares elétron-pósitron pela interação com a luz de fundo extragaláctica (*Extragalactical Background Light* (EBL)): ($\gamma_{VHE} + \gamma_{EBL} \rightarrow e^+ + e^-$), que podem sofrer espalhamento compton e gerar raios gama em um processo de cascata.

A atenuação do fluxo observado (F_{obs}) em relação ao fluxo da fonte (F_{font}) é expressa como:

$$F_{obs}(E) = F_{font}(E)e^{-\tau(E,z)}, \quad (31)$$

onde $\tau(E, z)$ é a profundidade óptica, que depende da energia do fóton (E) e do *redshift* (z) da fonte.

Este processo atenua o fluxo de fontes distantes, criando um horizonte de raios gama (definido por Aharonian (2004) como a distância para a qual $\tau = 1$), quanto maior a energia, menor o caminho médio livre e maior a atenuação.

Fótons de energia extrema (~ 100 TeV) também interagem com a CMB, como a densidade desta é bem maior que a EBL (~ 400 fótons cm^3), raios gama com essa energia raramente chegam à Terra.

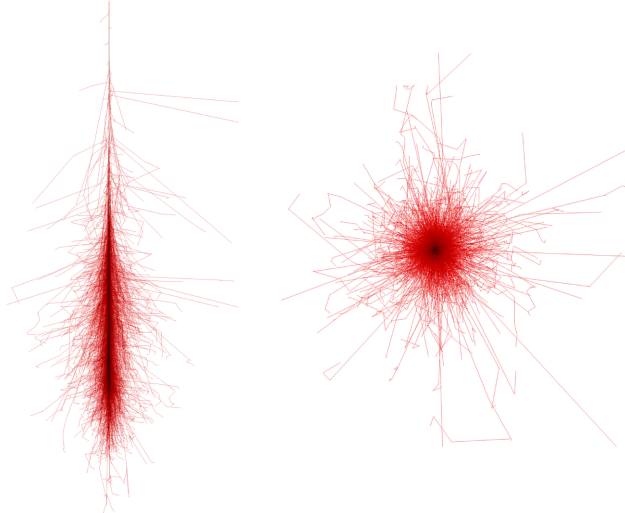
2.3 Chuveiros Atmosféricos Extensos

Essas partículas percorrem grandes distâncias no espaço, sendo aceleradas e sujeitas a perturbações até, eventualmente, alcançarem a atmosfera terrestre. No momento de sua chegada, encontram um meio opaco à sua propagação, colidindo com os átomos constituintes da atmosfera, predominantemente hidrogênio, oxigênio e nitrogênio (GILER, 2019). Tais interações geram uma cascata de partículas secundárias, conhecida como Chuveiro Atmosférico Extenso (EAS), que dá origem a múons, píons, neutrinos, pares elétron-pósitron, fótons, entre outras partículas secundárias mais leves e menos energéticas, que podem decair, sofrer novas colisões ou até mesmo atravessar o planeta praticamente sem interações (como no caso dos neutrinos) (ARBELETICHE, 2021) (SILVA, 2024).

A morfologia e a composição dessas cascatas dependem da natureza da partícula primária, como apontado em Silva (2024), essas cascatas se desenvolvem na atmosfera tanto longitudinalmente quanto lateralmente, criando uma cascata característica, enquanto um raio gama gera uma cascata puramente eletromagnética, tendo um perfil mais compacto (figura 23), um raio cósmico gera uma cascata mais complexa, tendo componentes eletromagnéticas, hadrônicas e muônicas, apresentando um perfil mais irregular (figura 24). Uma das formas de

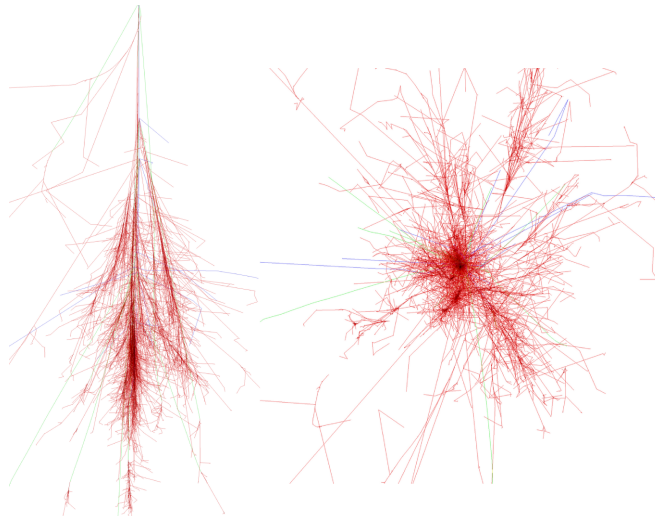
distinguir a partícula primária geradora da cascata é através dessas diferenças no desenvolvimento dos chuviros.

Figura 23 – Desenvolvimento longitudinal e lateral de EAS iniciado por um raio gama de 1 TeV



Fonte: Adaptado de Schmidt e Knapp (2005).

Figura 24 – Desenvolvimento longitudinal e lateral de EAS iniciado por um próton de 1 TeV



Fonte: Adaptado de Schmidt e Knapp (2005).

2.3.1 Chuviros eletromagnéticos

O desenvolvimento de cascatas iniciadas por fótons ou elétrons é puramente eletromagnético. O modelo que descreve o desenvolvimento de chuviros eletromagnéticos na atmosfera é o Modelo de Heitler. Neste modelo, assume-se que um fóton primário de energia E_0 percorre uma distância X_0 (Comprimento de Radiação $X_0 \sim 36,7 \text{ g cm}^{-2}$ no ar (SPURIO, 2018)) antes de ser convertido em um par $e^- + e^+$, dividindo a energia do raio gama de forma igual, essas partículas então percorrem X_0 e emitem um fóton. Esse processo em cascata continua até que

a energia média das partículas caia abaixo da energia crítica E_c de ≈ 86 MeV no ar (GRUPEN, 2020).

O número máximo de partículas secundárias (N_{max}) geradas pelo chuvaire depende diretamente da energia da partícula primária (E_0):

$$N_{max} = \frac{E_0}{E_c}. \quad (32)$$

Quanto mais energética for a partícula primária, maior é o número de partículas secundárias geradas por ela e maior o poder de penetração do chuvaire na atmosfera. A profundidade atmosférica onde o número máximo de partículas ocorre em chuvaires eletromagnéticos ($X_{max}^{(EM)}$) é obtida através da relação:

$$X_{max}^{(EM)} \approx X_0 \log_2(N_{max}). \quad (33)$$

2.3.2 Chuveiros hadrônicos

Os chuvaires iniciados por raios cósmicos (hádrons), são descritos pelo modelo de Heitler-Matthews: Ao colidir com um núcleo de ar, o próton primário produz mésons secundários. Cerca de $2/3$ das partículas geradas são píons carregados (π^\pm , cujo decaimento principal em $u^\pm + \nu_u$ injeta a componente muônica e de neutrinos no chuvaire) e $1/3$ são píons neutros (π^0), que rapidamente decaem em $\gamma\gamma$. A energia total após n interações é a soma da energia das componentes hadônicas e eletromagnéticas, dadas pelas fórmulas:

$$\begin{aligned} E_{had} &= \left(\frac{2}{3}\right)^n E_0 \\ E_{EM} &= \left[1 - \left(\frac{2}{3}\right)^n\right] E_0 \end{aligned} \quad (34)$$

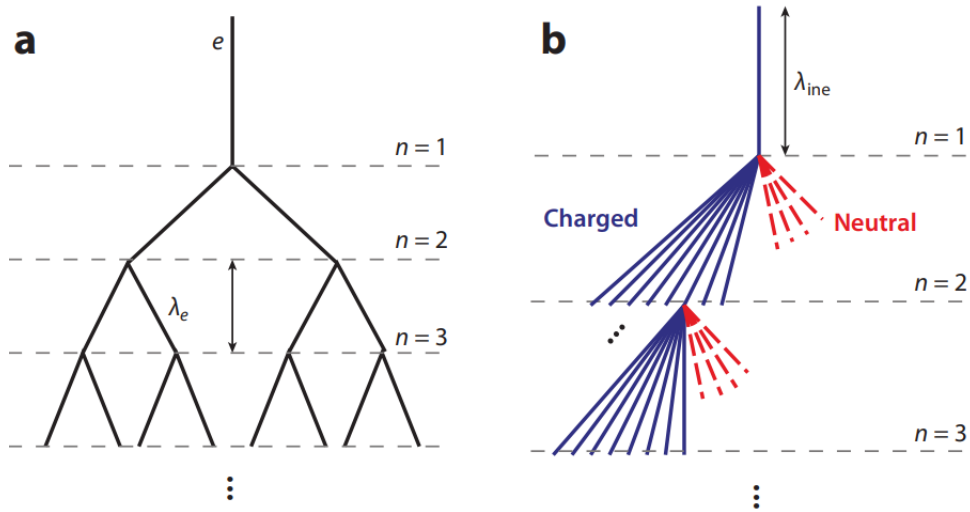
e o desenvolvimento máximo do chuvaire hadrônico ($X_{max}^{(had)}$) acontece após a produção de n_{total} partículas, sendo dado por:

$$X_{max}^{(had)} = \lambda_I + X_{max}^{(EM)} \left[\frac{E_0}{2n} \right]. \quad (35)$$

Sendo λ_I o comprimento de interação hadrônica, análogo ao X_0 visto anteriormente, é a distância média que um hádron percorre antes de colidir com um átomo na atmosfera. Para prótons no ar, esse valor é de ~ 90 g/cm² (SPURIO, 2018).

Diferente dos elétrons, os múons gerados em uma cascata hadrônica sofrem poucas perdas radiativas e de espalhamento na atmosfera, conseguindo alcançar o solo. A quantidade de múons produzidos depende da massa A da partícula primária, seguindo uma lei de potência

Figura 25 – a) Modelo de Heitler para cascatas eletromagnéticas e b) Modelo de Heitler-Matthews para cascatas hadrônicas

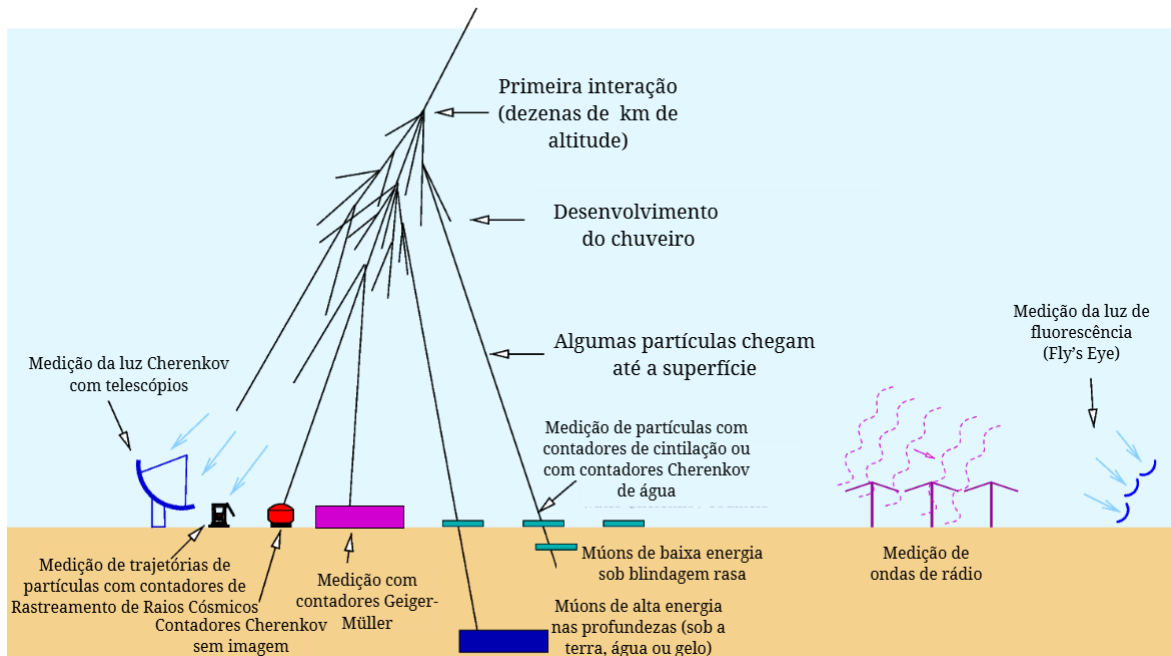


Fonte: Arbeletche (2021).

$N_\mu \propto A^{1-\beta} E_0^\beta$ (onde $\beta \approx 0,9$), sendo esta outra forma de distinguir entre chuviscos iniciados por prótons e por núcleos mais pesados, como apontam Giler (2019) e Arbeletche (2021).

Como a atmosfera impede que essas partículas atinjam o chão, uma forma de estudá-las é através da observação e análise de seus subprodutos (Figura 26).

Figura 26 – Diferentes técnicas utilizadas para a detecção de EAS



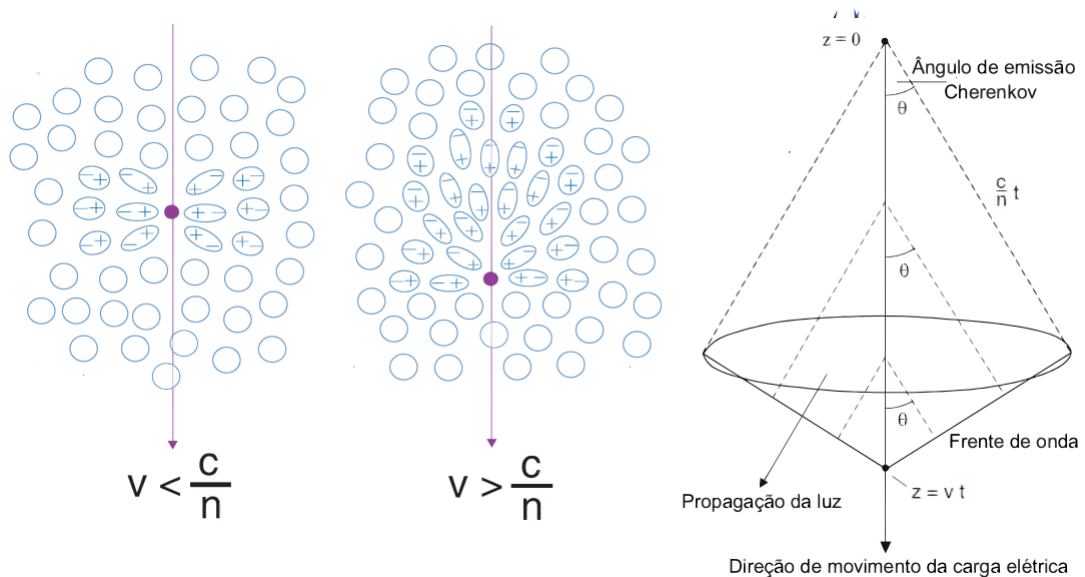
Fonte: Adaptado de Berniöhr (2007).

2.3.3 Emissão de Radiação Cherenkov em EAS

A Radiação Cherenkov é um fenômeno eletromagnético descoberto experimentalmente por Cherenkov (1934) e explicado posteriormente por Frank e Tamm (1937). Esse fenômeno ocorre quando uma partícula viaja com velocidade maior que a velocidade da luz em determinado meio e causa um *boom* fotônico análogo a *boom* sônico causado quando algo viaja mais rápido que o som. Após anos de estudos sobre raios cósmicos utilizando câmaras de nuvens, Blackett (1948) propôs que eles poderiam gerar radiação Cherenkov e contribuir para a luz de fundo do céu noturno.

O esquemático na Figura 27 mostra que, ao atravessar um meio dielétrico, como o ar, uma partícula carregada e com velocidade v , induz uma polarização nas moléculas vizinhas ao longo de sua trajetória. Se a velocidade da partícula excede a velocidade da luz no meio ($v > c/n$), ela cria uma onda de choque eletromagnética coerente que resulta na emissão de radiação Cherenkov (HORVATH, 2022; WEEKES, 2003)

Figura 27 – Polarização do meio e emissão de radiação Cherenkov



Fonte: Souto (2018).

Para que ocorra a emissão de radiação, é necessário que a partícula tenha uma energia mínima (E_{th}) que depende da massa de repouso da partícula (m_0).

$$E_{th} = \frac{m_0 c^2}{1 - n^{-2}}. \quad (36)$$

Se olharmos para o resultado dessa equação no nível do mar ($n \approx 1,00029$), Giler (2019) nos dá o limiar para algumas partículas que compõem o EAS: ~ 21 MeV para elétrons, $\sim 4,4$ GeV para múons, ~ 39 GeV para prótons e da ordem de TeV para núcleos de Ferro.

A quantidade de fótons Cherenkov emitidos por comprimento de onda pode ser obtida a partir da Fórmula de Frank-Tamm, chegando a:

$$\frac{d^2N}{d\lambda dx} = \frac{2\pi Z^2}{\lambda^2} \sin^2 \theta_C \propto \frac{1}{\lambda^2}. \quad (37)$$

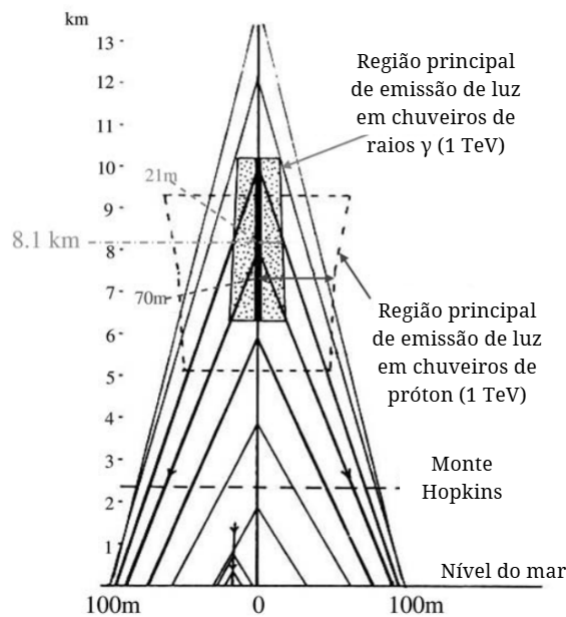
A partir dela é possível perceber teoricamente que a maior quantidade de fótons será emitida em regiões de menor comprimento de onda, isso explica a coloração azulada da radiação Cherenkov (HORVATH, 2022).

A radiação é emitida em um formato de cone de luz com Ângulo θ_C ao redor da trajetória da partícula, dado por:

$$\theta_C = \arccos\left(\frac{1}{n\beta}\right). \quad (38)$$

A atmosfera não é homogênea, isso faz com que o ângulo de emissão varie entre $\sim 0,7^\circ$ a uma altitude de 10 km até $\sim 1,4^\circ$ no nível do mar (figura 28). Resultando em uma poça de luz de formato anelar, também chamado de Anel Cherenkov, com tamanho entre 80 a 120 m na superfície e que contém cerca de 10^6 fótons Cherenkov (para um raio gama de 1 TeV), com duração e alguns nanossegundos.

Figura 28 – Ângulo de emissão de radiação Cherenkov em diferentes altitudes na atmosfera



Fonte: Adaptado de Weekes (2003).

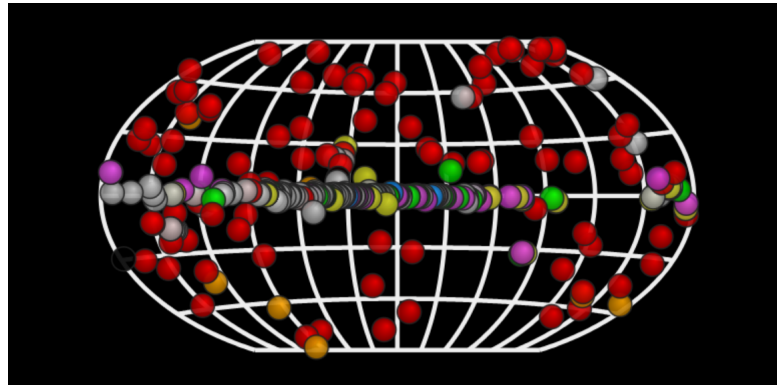
2.4 Imageamento Cherenkov Atmosférico

No final da década de 1950, Giuseppe Cocconi propôs que fontes astrofísicas como a Nebulosa do Caranguejo, deveriam emitir raios gama de alta energia que poderiam ser detectados. Na década de 1960, na União Soviética, a equipe de Alexander Chudakov construiu o primeiro grande arranjo de telescópios para fontes pontuais de raios gama, marcando o nascimento prático da astronomia Cherenkov (MIRZOYAN, 2022).

Os avanços começaram a aparecer com a construção dos telescópios de segunda geração, com destaque para o Whipple, usado para detectar o sinal de raios gama da Nebulosa do Caranguejo por Weekes *et al.* (1989).

A terceira geração de IACTs começa a ser pensada no final da década de 1990 e construída no início dos anos 2000. Constituída pelos observatórios *High Energy Stereoscopic System* (HESS), na Namíbia, *Very Energetic Radiation Imaging Telescope Array System* (VERITAS), nos EUA, e *Major Atmospheric Gamma Imaging Cherenkov* (MAGIC), na Espanha. O grande diferencial dessa geração foi a implementação da observação estereoscópica, com vários telescópios para observar o mesmo evento, permitindo a reconstrução tridimensional do chuveiro e melhorando significativamente a qualidade dos dados obtidos e elevando o número de fontes conhecidas de algumas dezenas para mais de 320 (Figura 29) .

Figura 29 – Atualmente existem mais de 320 fontes catalogadas, entre galáticas, extragaláticas e desconhecidas. Informações obtidas do site TeVCat em março de 2026



Fonte: Wakely e Horan (2008).

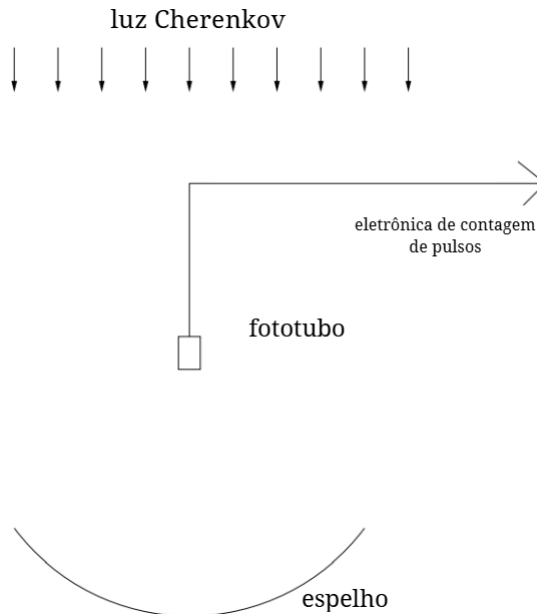
2.4.1 Princípios de funcionamento

Como a densidade de fótons Cherenkov no solo é baixa, são necessários grandes refletores ópticos, com centenas de m^2 . O espelho principal coleta a luz e a focaliza em uma câmera posicionada no plano focal, a figura 30 ilustra um esquema básico de um IACT. Embora necessite de espelhos grandes, a ótica não precisa ter a mesma precisão construtiva dos telescópios de astronomia óptica, tornando sua construção mais barata.

A câmera é composta por uma matriz de fotodetectores de alta sensibilidade e resposta rápida, historicamente Tubos Fotomultiplicadores (*Photo Multiplier Tube* (PMT)) e, mais recentemente, fotomultiplicadores de silício (*Silicon Photomultiplier* (SiPM)) (AMBROSI *et al.*, 2013; CORTINA, 2019). A eletrônica de aquisição deve ser super-rápida, visto que a duração do *flash* de radiação Cherenkov é de alguns nanossegundos.

A imagem formada na câmera é uma projeção bidimensional do desenvolvimento tridimensional do chuveiro. A morfologia dessa imagem é a chave para a separação entre Raios Gama (mais compactos e simétrico ao redor da trajetória da partícula primária) e Raios Cós-

Figura 30 – Esquema simplificado do sistema de detecção de um telescópio Cherenkov.



Fonte: Weekes (2003).

micos (que resultam em imagens mais irregulares), a figura 31 mostra a distribuição lateral, temporal e angular típica de um chuveiro hadrônico e de um chuveiro gama.

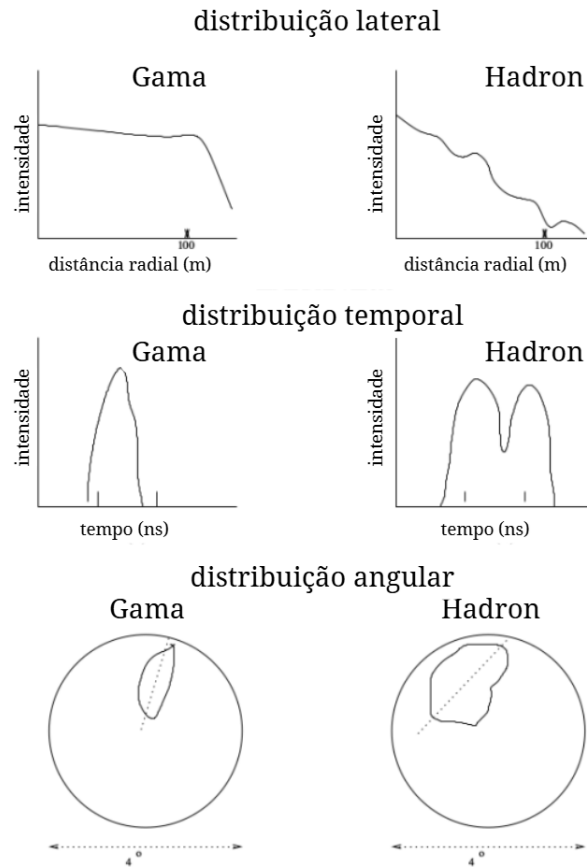
Raios Gama ainda podem ser divididos em Raios Gama Pontuais e Raios Gama Difusos, neste trabalho tratamos como pontuais aqueles no qual a origem no céu aponta para a fonte que está sendo observada, por outro lado, raios gama difusos são aqueles que atingem o sensor, mas cujas direções de reconstrução situam-se fora da região ocupada pela fonte observada.

A origem dos raios gama difusos pode estar ligada a fontes astrofísicas localizadas nas proximidades do objeto principal, à própria radiação difusa da galáxia e a componente eletromagnética de chuveiros hadrônicos. Tipicamente, raios gama difusos são tratados como ruídos e separados dos raios gama pontuais (sinal). Visando as novas estratégias observacionais previstas pelo CTAO, como a realização de mapeamentos de céu profundo e o uso de apontamento divergente, os resultados apresentados neste trabalho não fazem distinção entre gama difuso e gama pontual, dado que a capacidade de classificar e reconstruir eventos eletromagnéticos em regiões do campo de visão fora do alvo pode representar uma estratégia para a otimização do instrumento durante as futuras operações do observatório.

2.4.2 Cherenkov Telescope Array Observatory

O CTAO é o expoente da próxima geração de observatórios terrestres de raios gama, projetado para melhorar a sensibilidade em uma ordem de grandeza em relação aos instrumentos atuais de 3ª geração e ampliar a cobertura de energia para a faixa entre 20 GeV a 300 TeV (Figura 32) (ACHARYA *et al.*, 2018).

Figura 31 – Distribuição lateral, temporal e angular dos fotoelétrons capturados pela câmera em um IACT



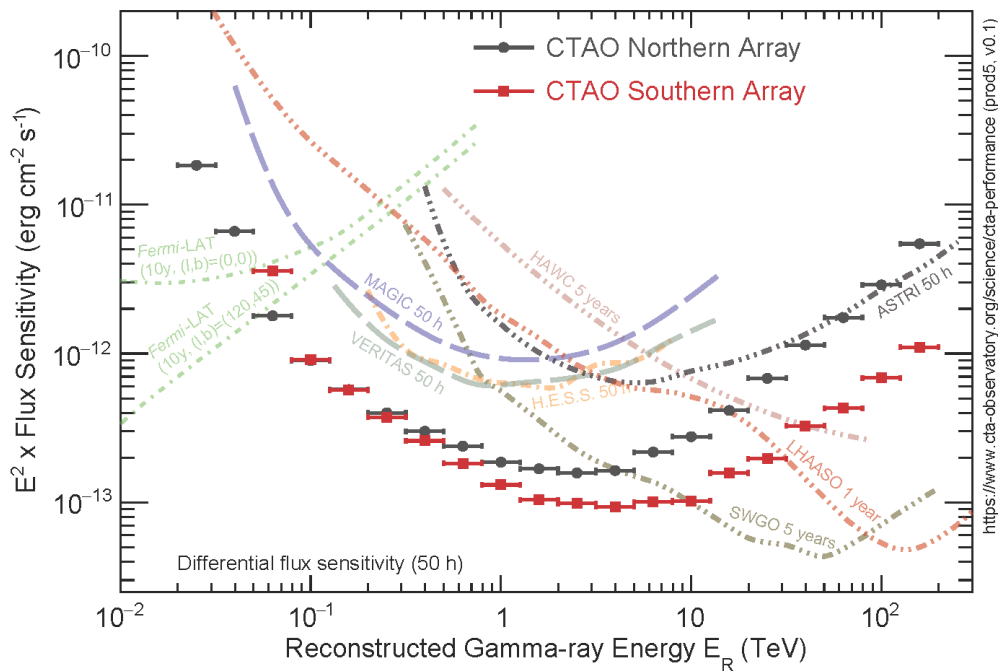
Fonte: Weekes (2003).

A ampla faixa de energia do CTAO se dá pela utilização de telescópios de diferentes tamanhos, como mostra a figura 33 .

1. *Small-Sized Telescope (SST)*: para energias mais altas ($> 5 \text{ TeV}$), possui cerca de 4 m de diâmetro. Energias mais altas produzem mais luz, não sendo necessário uma superfície coletora tão grande, porém são eventos raros e é necessário cobrir uma área de solo muito extensa para maximizar as chances de detectá-los.
2. *Medium-Sized Telescope (MST)*: De tamanho intermediário, conta com espelhos de 12 m e cobre a faixa central (de 150 GeV a 5 TeV). Equilibra área de coleta e campo de visão.
3. *Large-Sized Telescope (LST)*: O maior dos telescópios do CTAO, possui 23 m de diâmetro e foi otimizado para chuvaros de baixas energias, com foco na astrofísica extragalática, investigará AGNs em *redshift* $z \leq 3$, GRBs em $z \leq 10$ e, devido à sua capacidade de reapontamento rápido, é ideal para observar eventos transitórios.

Possui um sistema de óptica ativa, com 198 espelhos hexagonais de 1,5 m que totaliza $\sim 396 \text{ m}^2$ de área refletora. Cada espelho possui 3 pontos de fixação, sendo um afixado

Figura 32 – Sensibilidade projetada para o CTAO com 50h de observação em comparação com outros instrumentos



Fonte: CTA Observatory (2021).

e outros dois ligados a atuadores para correção do posicionamento do espelho, esse sistema alcança resolução de $5\mu\text{m}$.

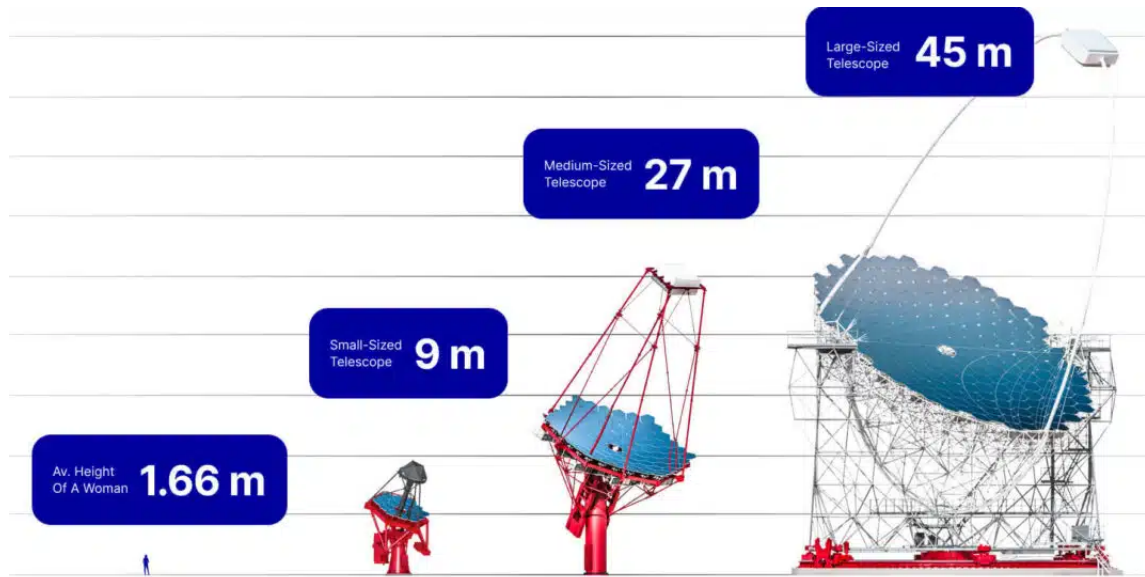
O imageamento da luz Cherenkov é feito por uma câmera de ~ 3 m de diâmetro composta por 1855 PMTs de alta eficiência quântica acoplados a tubos de Winston que focalizam a luz recebida (CORTINA, 2019). Os PMTs são agrupados de 7 em 7 em *clusters* (figura 34) contendo um sistema eletrônico de leitura ultra rápida baseada em chips DSR4, que alcançam uma taxa de leitura de até 5 GHz (YANG *et al.*, 2015).

Embora o projeto original previsse a instalação de mais de 100 telescópios, restrições orçamentárias levaram a uma revisão do projeto, que agora prevê a instalação de 64 telescópios (Configuração Alfa) divididos em dois sítios, um em cada hemisfério, para garantir a cobertura completa do céu (HOFMANN; ZANIN, 2024).

2.4.2.1 CTAO - Norte

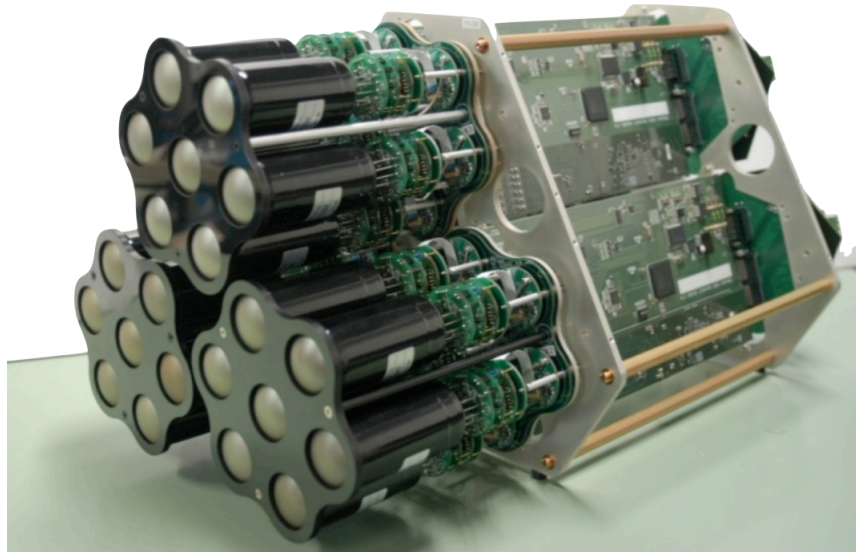
Localizado no observatório Roque de Los Muchachos, em La Palma - Espanha. Será composto por 4 LSTs e 9 MSTs, cobrindo uma área de $0,25 \text{ km}^2$. Sua otimização é para energias mais baixas, focando no estudo de eventos transientes e objetos astrofísicos distantes.

Figura 33 – Comparação em escala mostrando a altura das três classes de telescópios projetadas para o CTAO



Fonte: CTA Observatory (2024a).

Figura 34 – Cluster de PMTs, contendo 21 pixels cada, acoplados a um sistema eletrônico de leitura ultra-rápida

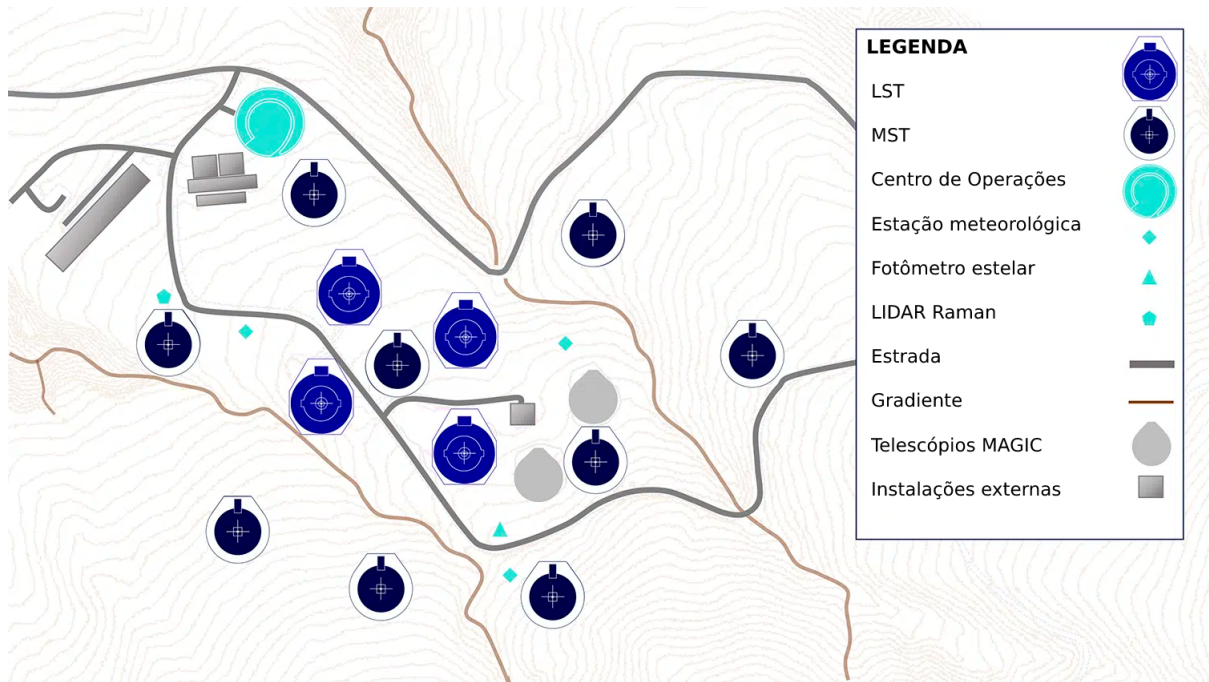


Fonte: Ambrosi *et al.* (2013).

2.4.2.2 CTAO - Sul

Localizado no Cerro Paranal, no Deserto de Atacama - Chile, com uma vista privilegiada para o centro da Galáxia. Conta com 14 MSTs e 37 SSTs cobrindo uma área de $\sim 3 \text{ km}^2$. Fará o imageamento do plano galáctico em busca de PeVatrons, e buscará sinais de matéria escura no centro da Galáxia. Otimizado para energias entre 150 GeV e 300 TeV.

Figura 35 – Distribuição geográfica dos telescópios da configuração alfa no sítio CTAO Norte, além dos 4 LSTs e dos 9 MSTs, a imagem também mostra os 2 telescópios MAGIC



Fonte: CTA Observatory (2024a).

2.5 Métodos Tradicionais de Reconstrução

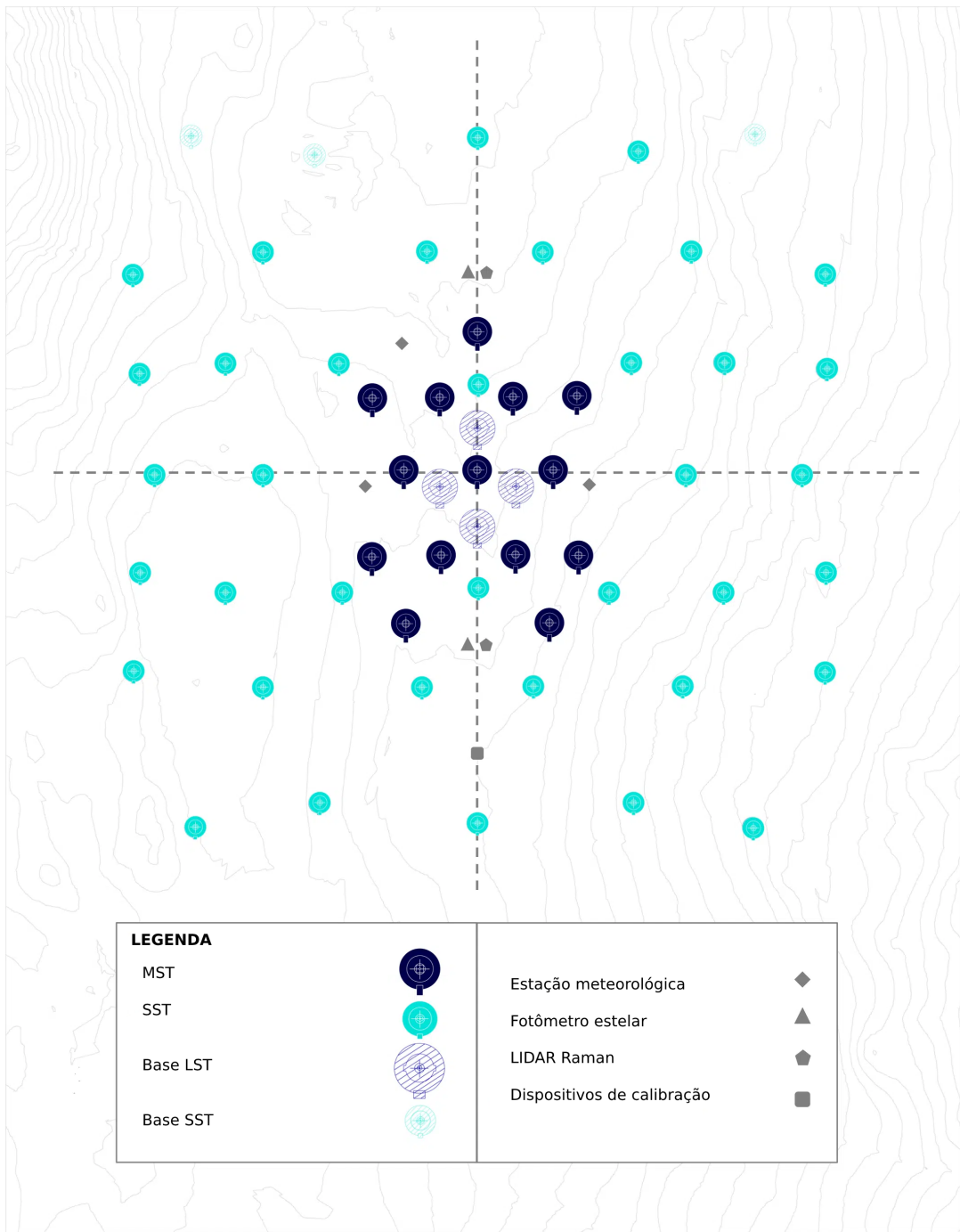
A reconstrução das propriedades da partícula primária (tipo, energia e direção de chegada) a partir das imagens Cherenkov é um problema inverso complexo. Os métodos desenvolvidos para a análise podem ser divididos em duas abordagens: a abordagem paramétrica e a abordagem baseada em modelos.

2.5.0.1 Abordagem paramétrica - Parâmetros de Hillas

A abordagem paramétrica é a técnica mais tradicional e mais difundida na astronomia de raios gama. HILLAS (1985) propôs que a imagem da luz Cherenkov obtida pela câmera pode ser aproximada por uma elipse a partir da qual se faz o uso de parâmetros baseados na largura, no comprimento e na orientação dessa elipse. Com base nesses parâmetros é possível estimar a direção, a energia e do ponto de impacto da partícula primária, sendo possível também fazer a distinção entre chuviscos iniciados por Raios Gama daqueles iniciados por Raios Cósmicos (ARBELETICHE, 2021).

A figura 37 ilustra os parâmetros geométricos derivados da distribuição de luz no plano focal do telescópio (sistema de coordenadas X-Y), modelada como uma elipse. O ponto C representa o centroide da imagem, enquanto M representa o centro do campo de visão do telescópio. Os parâmetros morfológicos principais, *Length* (comprimento) e *Width* (largura), correspondem, respectivamente, à dispersão longitudinal e lateral do sinal ao longo dos eixos

Figura 36 – Distribuição geográfica dos telescópios da configuração alfa no sítio CTAO Sul, além dos 14 MSTs e dos 37 SSTs, a imagem também mostra os locais de fundação para futura expansão com 3 SSTs e 4 LSTs.

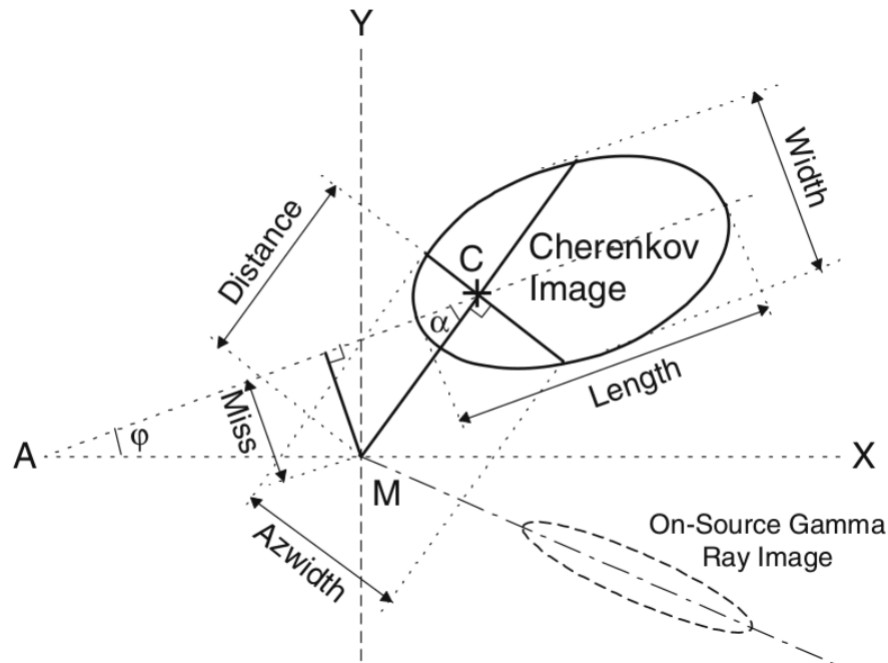


Fonte: CTA Observatory (2024a).

maior e menor da elipse. A separação angular entre o centroide C e a origem M define o parâmetro $Distance$, enquanto o $Miss$ representa a distância perpendicular entre o prolongamento do eixo maior da elipse e o centro da câmera. O parâmetro $Azwidth$ quantifica a distribuição de luz perpendicular à linha que une o centroide à origem (GILER, 2019).

A combinação destes parâmetros é frequentemente utilizada em métodos de aprendizado supervisionado, como *Random Forests* ou *Boosted Decision Trees*, mantendo uma efici-

Figura 37 – Parametrização de Hillas de uma imagem obtida por IACTs



Fonte: Giler (2019).

ência razoável para os raios gama (NIETO *et al.*, 2017; SHARMA *et al.*, 2014). Contudo, essa abordagem implica uma perda de informação contida na correlação entre pixels individuais, o que motiva o uso de técnicas de *Deep Learning* que operam diretamente sobre as imagens.

2.5.0.2 Abordagem Baseada em Modelos

O método ImPACT (*Image Pixel-wise fit for Atmospheric Cherenkov Telescopes*) proposto por Parsons e Hinton (2014) consiste na criação de uma biblioteca de templates simulados usando Monte Carlo. Em vez de reduzir a imagem capturada a elipses estatísticas, o método preserva a informação de intensidade de cada um dos pixels da câmera, tratando o chuveiro como uma distribuição contínua de luz.

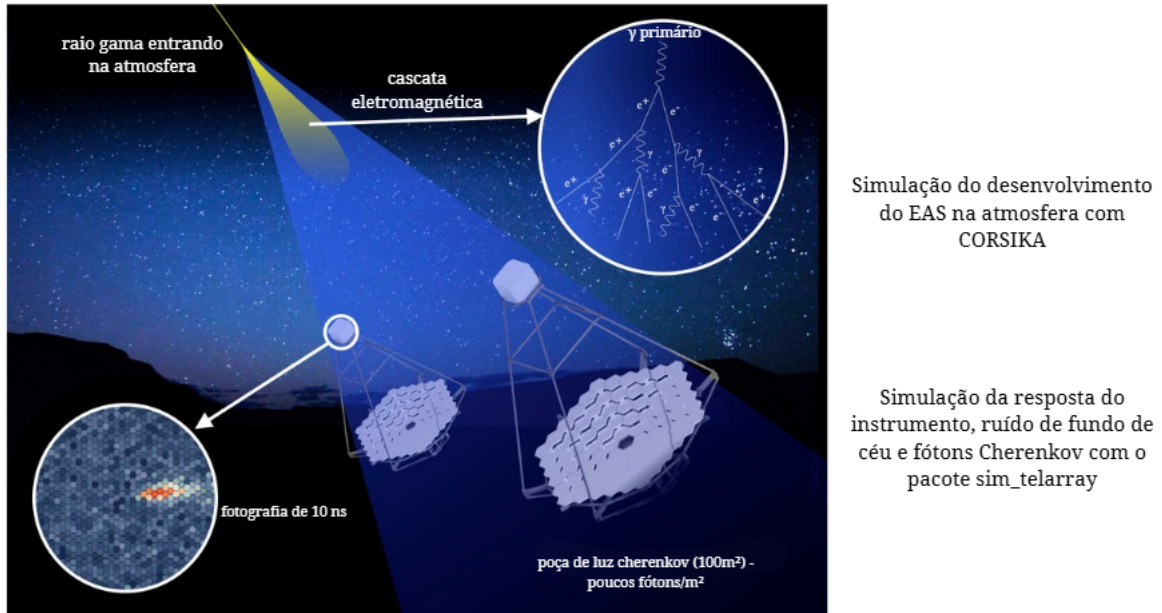
A reconstrução do chuveiro é feita através de análise de verossimilhança comparando pixel a pixel a imagem real detectada pelo telescópio com os templates pré-calculados na biblioteca. Essa abordagem maximiza a extração de informação morfológica do chuveiro, resultando em uma melhoria significativa na sensibilidade do observatório e na capacidade de rejeição de ruído.

2.6 Simulações Monte Carlo

Diante da impossibilidade a criação de chuveiros atmosféricos controlados para calibrar e validar o funcionamento dos IACTs, a análise dos dados depende da comparação com simu-

lações computacionais que reproduzem a física do desenvolvimento da cascata na atmosfera. A simulação utilizada para IACTs consiste em simular o desenvolvimento do chuvaire e a resposta do detector (Figura 38, para a primeira tarefa, o programa mais utilizado é o CORSIKA, e para a segunda, o `sim_telarray` (SILVA, 2024; SPURIO, 2018).

Figura 38 – Simulações IACT com CORSIKA e `sim_telarray`



Fonte: Adaptado de (WHITE; BERNLOHR; DESY, 2018).

2.6.1 CORSIKA

O *COsmic Ray Simulations for KAscade* (CORSIKA). Originalmente desenvolvido para o experimento KASCADE em Karlsruhe, Alemanha, é um programa de simulação de Monte Carlo que descreve a evolução de EAS na atmosfera iniciados por fótons, prótons, núcleos ou outras partículas primárias. O código rastreia as partículas secundárias individualmente ao longo de suas trajetórias na atmosfera, computando probabilisticamente suas interações com os núcleos atômicos do ar, os decaimentos de partículas instáveis, as perdas de energia por ionização e as deflexões causadas por espalhamento (HECK *et al.*, 1998).

A primeira versão foi desenvolvida na década de 1980, baseada em 3 simuladores utilizados na época: O conjunto de programas desenvolvidos por Grieder forneceu a estrutura geral do programa adotada no CORSIKA. Além disso, suas rotinas do modelo ISOBAR foram incorporadas para servir como uma alternativa para a simulação de interações hadrônicas em baixas energias. O Gerador de Interações HDPM (*Hybrid Dual Parton Model*), desenvolvido por Capdevielle, que tinha como função descrever as interações hadrônicas de prótons em altas energias, ajustando-se bem aos dados de colisores disponíveis na época. E o EGS4 (*Electron Gamma Shower system*), utilizado na simulação de detectores em experimentos de física de partícu-

las, foi incorporado ao CORSIKA para simular a parte eletromagnética do chuvaire atmosférico (HECK *et al.*, 1998).

Embora o CORSIKA simule as interações físicas, a geração da radiação Cherenkov requer o uso do pacote desenvolvido por Konrad Bernlöhrr, conhecido como pacote IACT/Bernlöhrr. A compilação do CORSIKA com a opção IACT introduz algumas modificações no processo de simulação, a principal delas diz respeito à geometria de detecção, considerando os detectores como esferas na superfície, apenas os fótons que atravessam essas esferas são efetivamente salvos no arquivo, reduzindo o volume de dados.

2.6.2 sim_telarray

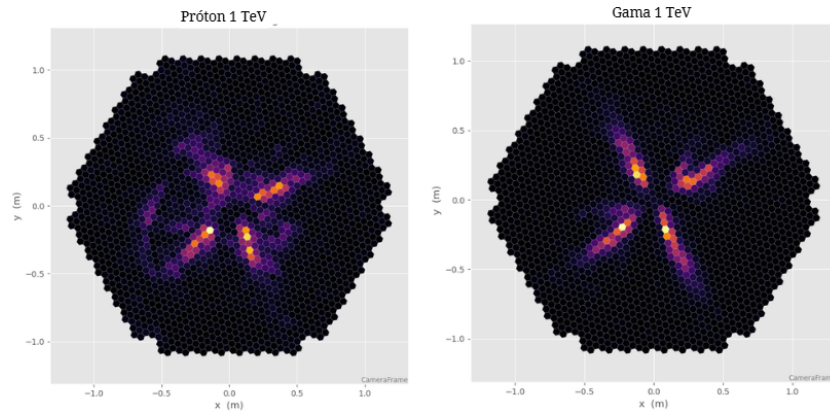
Após a geração dos chuvaire atmosféricos e da produção da radiação Cherenkov pelo CORSIKA, a próxima etapa consiste na simulação da resposta do detector. Para isso usa-se do pacote sim_telarray (*SIMulation of TELEscope ARRAY*), desenvolvido por Konrad Bernlöhrr, originalmente para o experimento HEGRA e posteriormente adaptado para o H.E.S.S. e para o CTAO. O sim_telarray é projetado para simular a resposta de arranjos de IACTs, permitindo a configuração de múltiplos telescópios posicionados em locais arbitrários, onde cada telescópio pode possuir suas respectivas características ópticas e eletrônicas (BERNLÖHR, 2008).

O software lê os pacotes de fótons gerados pelo CORSIKA e realiza o rastreamento detalhado deles até a digitalização do sinal. O processo pode ser dividido em três estágios:

1. *Ray-Tracing*: A transmissão final da luz Cherenkov até o detector é refinada no sim_telarray, simulando a absorção por ozônio e o espalhamento Mie e Rayleigh em função do comprimento de onda e da altitude. Ao atingir a estrutura do telescópio, o software executa um *ray-tracing* completo. Ele considera a geometria segmentada e a refletividade dos espelhos, e os efeitos de sombreamento causados pela estrutura de sustentação da câmara (BERNLÖHR, 2008).
2. Resposta do detector: Os fótons que atingem o plano focal interagem com os sensores. O sim_telarray simula a eficiência quântica dos fotodetectores e a eficiência de coleta dos cones de Winston. Além de adicionar o ruído de fundo, simulando fótons aleatórios que atingem a câmara (BERNLÖHR, 2008).
3. Digitalização: O sinal elétrico gerado pelo sensor é processado. O software simula a formação do pulso eletrônico, a adição de ruído eletrônico e a digitalização feita pela eletrônica ultra-rápida (BERNLÖHR, 2008).

O resultado final da simulação é gravado em um arquivo idêntico ao arquivo gerado por um IACT real, do qual é possível extrair imagens como as da Figura 39 .

Figura 39 – Imagens obtidas em simulação de Monte Carlo

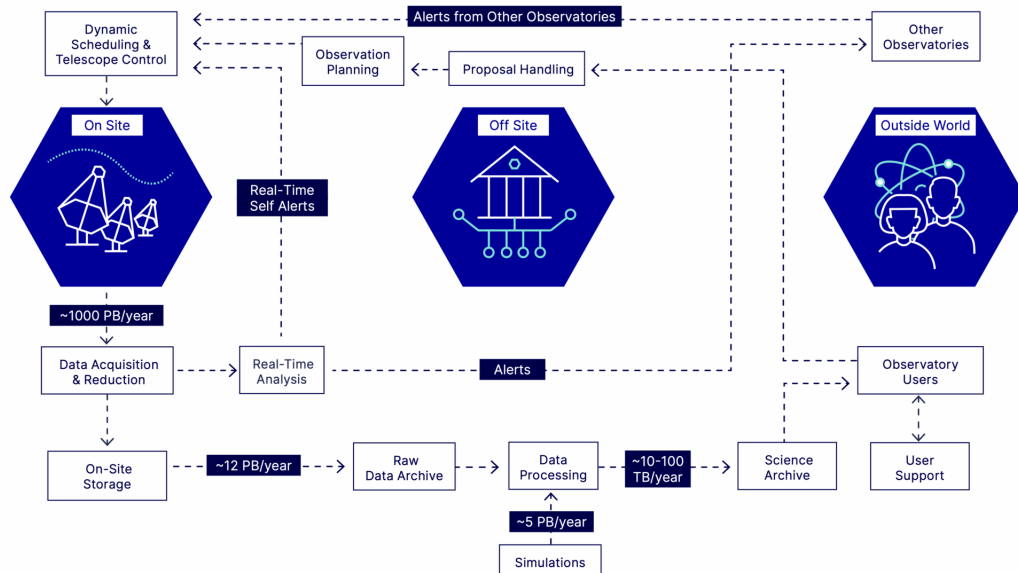


Fonte: Autoria própria.

2.7 Redes Neurais Artificiais e Aprendizado Profundo

A astronomia e a astrofísica de partículas encontram-se em meio a uma revolução impulsionada pelo avanço na instrumentação e nas técnicas de análise. O volume de dados gerados pelos observatórios modernos cresceu exponencialmente, a estimativa é de que o CTAO gere cerca de 2 PB de dados por noite (figura 40), constituindo um desafio que ultrapassa a capacidade de processamento tradicional.

Figura 40 – Fluxo de dados do CTAO após observações, os dados brutos (DL0) passam por etapas de processamento antes de serem disponibilizados ao público (DL3)



Fonte: Cherenkov Telescope Array Observatory ().

Em um cenário como este, a aplicação de técnicas de Aprendizado de Máquina (*Machine Learning* (ML)) torna-se praticamente obrigatória. NORRIS (2021) descreve essa transição como o surgimento da "*Large-n astronomy*", onde o foco dos estudos está na análise es-

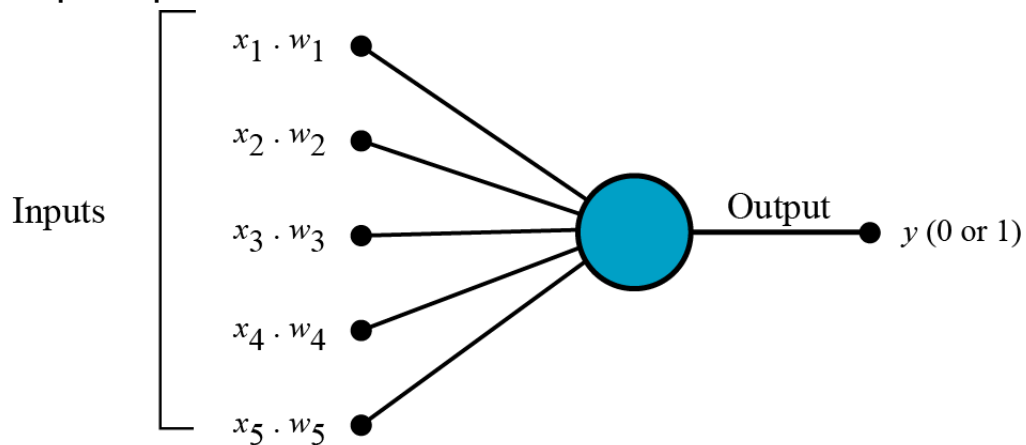
tatística de grandes populações de eventos. Para instrumentos como o LST-1 e o futuro CTAO, isso implica na necessidade de classificar e reconstruir propriedades de milhões de eventos de chuviros atmosféricos de forma rápida e automatizada.

As Redes Neurais Convolucionais (*Convolutional Neural Networks - CNNs*), originalmente desenvolvidas para visão computacional, demonstraram ser ideais para a classificação de fontes e eventos em astronomia (PARSONS; HINTON, 2014; NIETO *et al.*, 2017; SHILON *et al.*, 2019; SILVA; MELLO, 2025), permitindo a extração de características diretamente dos dados brutos ou minimamente processados.

2.7.1 Fundamentos de Redes Neurais

A unidade fundamental de uma *Rede Neural Artificial (RNA)* é o neurônio, uma abstração matemática inspirada no funcionamento do neurônio biológico. Um neurônio artificial realiza uma soma ponderada de suas entradas, seguida pela aplicação de uma função de ativação.

Figura 41 – Desenho esquemático de um perceptron, contendo entradas que são multiplicadas por um peso e resultam em uma saída



Fonte: Adaptado de Nielsen (2015).

O modelo mais simples é o Perceptron, mostrado na figura 41, foi proposto por Frank Rosenblatt em 1958. Matematicamente, o *Perceptron* pode ser representado pela Equação 39:

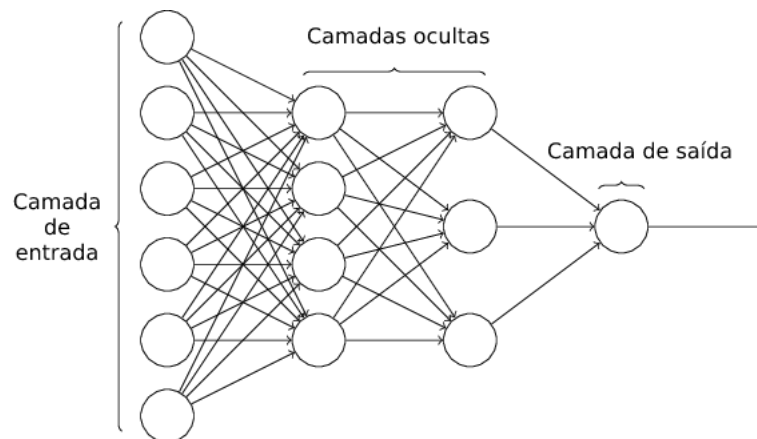
$$\text{saída} = \begin{cases} 0 & \text{se } \sum_j w_j x_j \leq \text{limiar} \\ 1 & \text{se } \sum_j w_j x_j > \text{limiar} \end{cases}, \quad (39)$$

ou ainda, uma modelagem mais generalizada para qualquer função de ativação f e um viés b :

$$\text{saida} = f \left(\sum_j w_j x_j + b \right). \quad (40)$$

Embora pioneiro, o Perceptron de camada única não tem capacidade de resolver problemas mais complexos. A superação dessa barreira ocorreu com o desenvolvimento do Perceptron Multicamadas (*Multi Layer Perceptron (MLP)*). Uma MLP, mostrada na figura 42, é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída.

Figura 42 – Esquemático de um MLP, camadas mais profundas extraem informações mais complexas



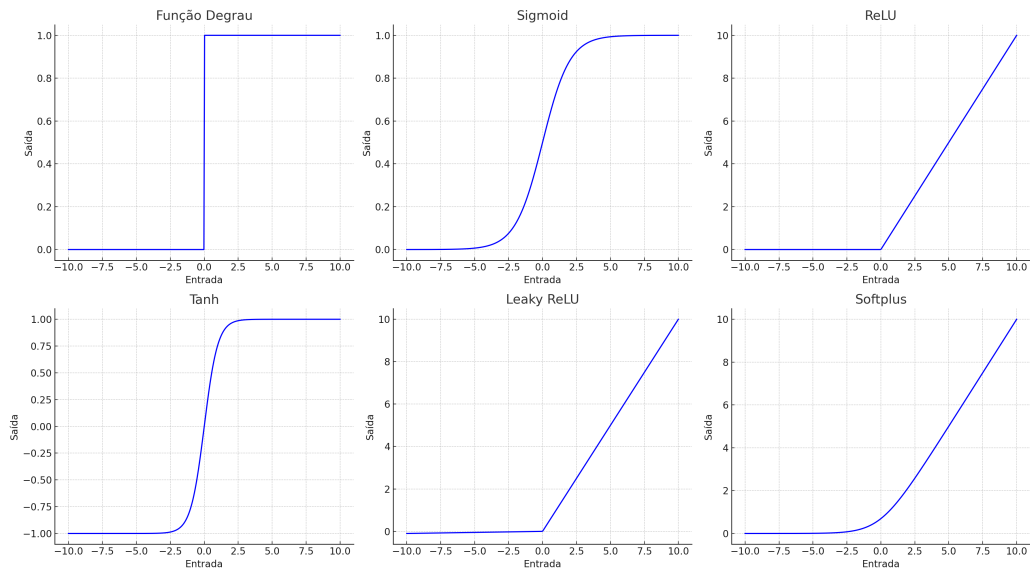
Fonte: Adaptado de Nielsen (2015).

Essa topologia permite que os computadores consigam "aprender" conceitos cada vez mais complexos, daí surge o conceito de aprendizado de máquina, que é "um sistema [...] treinado em vez de explicitamente programado. É apresentado exemplos relevantes para realizar uma tarefa e então o sistema encontrará uma estrutura estatística nesses eventos" (CORREIA; NOGUEIRA, 2023, p. 16). Sendo assim, tarefas repetitivas e relativamente complexas que envolvam, principalmente, o reconhecimento de padrões, podem ser automatizadas através do aprendizado de máquina.

2.7.1.1 Função de ativação

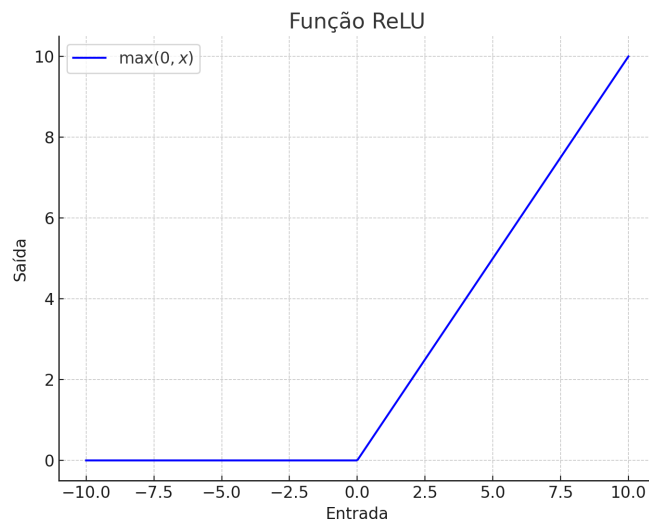
Também chamada de função de transferência, são fundamentais para o processo de aprendizado de máquina, são responsáveis por inserir a não-linearidade no aprendizado, a função de ativação decide "se a informação que o neurônio está recebendo é relevante para a informação fornecida ou deve ser ignorada"(GOODFELLOW; BENGIO; COURVILLE, 2016). A inserção da não-linearidade é o que permite que a RNA seja capaz de resolver problemas complexos. Existem diversas funções de ativação, como mostrado na Figura 43. Prince (2023) alega que a maior parte das redes neurais modernas usa como padrão a função de ativação ReLU, neste trabalho, além da ReLU, também foi utilizada a função Sigmoid.

Figura 43 – Diferentes funções de ativação utilizadas em redes neurais artificiais



Fonte: Silva (2024).

Figura 44 – Função Rectified Linear Unit - ReLU



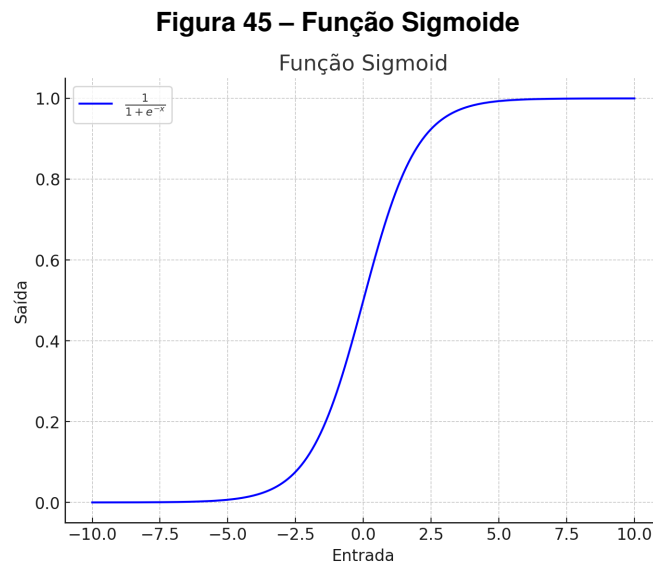
Fonte: Silva (2024).

2.7.1.1.1 ReLU

A função de ativação ReLU (Figura 44) é uma função não-linear. É amplamente utilizada em modelos de aprendizado profundo (DL) por conta de sua eficiência computacional e eficácia no treinamento de RNAs. Ela é definida matematicamente como o valor máximo entre zero e a entrada do neurônio:

$$f(x) = \max(0, x). \quad (41)$$

2.7.1.1.2 Sigmoide



A função sigmoide, ou logística (figura 45), também tem como característica a não-linearidade. Costuma ser bastante utilizada nas camadas de saídas em RNAs especialistas em classificação binária. Apresenta uma curva mais suave do que a ReLU, fazendo com que qualquer valor de entrada assuma valores que podem ser tratados como probabilidade de pertencer a uma classe. É representada matematicamente por:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (42)$$

2.7.2 Treinamento

O processo de "aprendizado" de uma rede neural se dá através de um treinamento, que nada mais é do que um problema de otimização onde se busca minimizar uma função de perda (*Loss Function*), através da alteração de um conjunto de parâmetros, quantificando o erro entre a saída da rede e o rótulo verdadeiro (*ground truth*) (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.7.2.0.1 Função de perda

A escolha da função de perda (ou *Loss function*) L é ditada pelo princípio da Estimativa de Máxima Verossimilhança. No contexto de DL, minimizar a função de perda é equivalente a maximizar a probabilidade logarítmica dos dados de treinamento sob o modelo estatístico definido pela rede (GOODFELLOW; BENGIO; COURVILLE, 2016).

Para tarefas de classificação binária, a função de perda padrão é a Entropia Cruzada Binária (*Binary Cross-Entropy* (BCE)). Nela, previsões p confiantes, mas erradas, são severamente punidas, o que força a rede a ajustar os pesos rapidamente no início do treinamento para que o resultado da previsão esteja de acordo com os rótulos dos dados de treinamento y . A BCE para uma amostra é definida como:

$$L_{BCE} = -[y \log(p) + (1 - y) \log(1 - p)]. \quad (43)$$

Em cenários onde o conjunto de dados apresenta classes desbalanceadas, como no caso deste trabalho, a BCE apresenta algumas limitações, a função de perda é dominada pelos eventos mais comuns, suprimindo o aprendizado dos eventos raros e difíceis. Uma alternativa à BCE para esses casos é apresentada por Yun *et al.* (2018): a *Focal Loss*. Uma função de perda baseada em BCE, com a adição de um fator modulador que atribui maior peso aos eventos mais difíceis:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t). \quad (44)$$

Onde p_t é a probabilidade prevista para a classe verdadeira, α_t é o fator de balanceamento e γ é o parâmetro de foco. Isso impede que a rede aprenda e fique viciada em eventos fáceis e força a rede a focar o aprendizado nos eventos limites.

Para as tarefas de regressão assume-se que o erro segue uma distribuição Gaussiana, a partir daí a maximização da verossimilhança leva à minimização do Erro Quadrático Médio (*Mean Squared Error* (MSE)).

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - p_i)^2. \quad (45)$$

Mas avaliar a perda total de uma rede que realiza tarefas distintas é um desafio, a rede precisa encontrar um valor que minimize tanto a tarefa de classificação quanto as tarefas de regressão de energia e direção. Uma função de perda global baseada na simples soma dos pesos pode fazer com que a função de perda de uma tarefa mais fácil seja preterida em função de outra mais difícil.

Em vez de definir hiperparâmetros estáticos para os pesos de cada tarefa, Kendall, Gal e Cipolla (2017) propõe a utilização de uma função composta baseada em incerteza homocedástica para a ponderação da perda. A maximização da log-verossimilhança leva a um objetivo de minimização onde a variância do ruído da tarefa (σ^2) atua como um fator de ponderação aprendível. Definindo $s = \log \sigma^2$, temos:

$$L = \sum_i \frac{1}{2} e^{-s_i} L_i + \frac{1}{2} s_i + \sum_j \frac{1}{2} e^{-s_j} L_j + \frac{1}{2} s_j. \quad (46)$$

Onde L_i é a função de perda da tarefa 1 e L_j é a função de perda da tarefa 2. Esse processo permite que o modelo dinamicamente atenue o gradiente de tarefas ruidosas ou difíceis (aumentando a incerteza s) durante as fases iniciais do treinamento, evitando o colapso da otimização.

2.7.2.0.2 Retropropagação ou otimização gradiente

Para minimizar a função de perda, procuramos parâmetros que façam com que o valor da predição seja o mais próximo da real possível, para isso são utilizados métodos baseados em gradiente, que inicializam os parâmetros de forma heurística e iteram sobre eles, ajustando-os a cada iteração.

Prince (2023) divide essa tarefa em dois passos: No primeiro passo computa-se o vetor gradiente da função de perda com base nos parâmetros iniciais ($\phi = [\phi_0, \phi_1, \dots, \phi_N]^T$), determinando a direção do movimento necessária para o refinamento dos pesos, representado matematicamente por:

$$\frac{\partial L}{\partial \phi} = \begin{bmatrix} \frac{\partial L}{\partial \phi_0} \\ \frac{\partial L}{\partial \phi_1} \\ \vdots \\ \frac{\partial L}{\partial \phi_N} \end{bmatrix}. \quad (47)$$

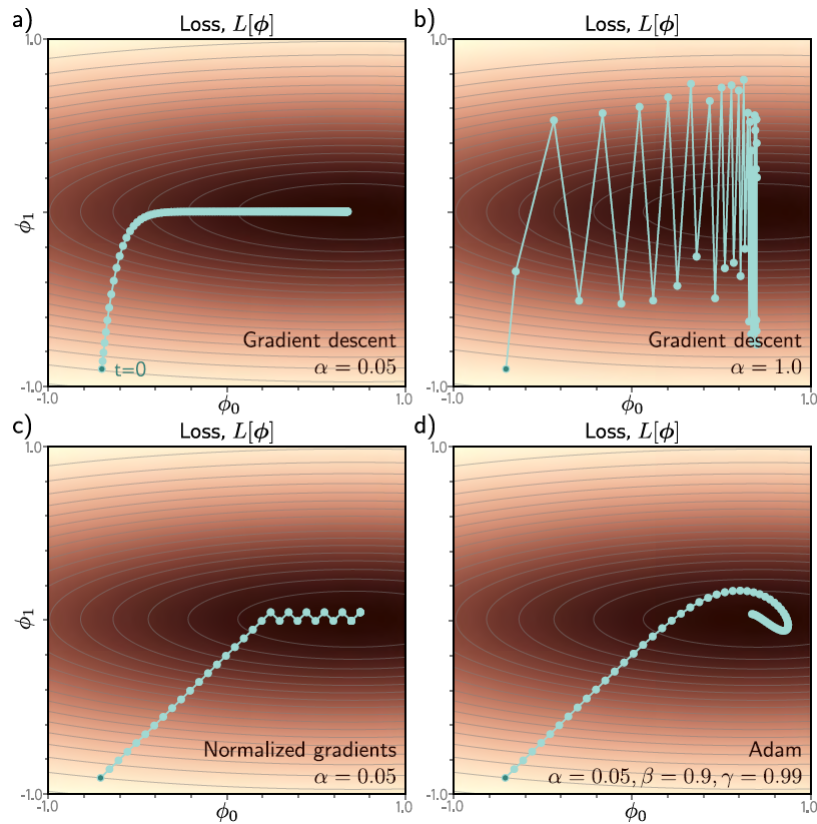
O segundo passo consiste na atualização dos parâmetros ϕ de forma que o valor computado se aproxime do valor de mínima da função de perda. Este ajuste é feito através da aplicação do hiperparâmetro α , chamado de taxa de aprendizagem (*learning rate*), que atua como um fator de escala para determinar o quanto deve-se "andar" em direção ao mínimo a cada iteração, é expresso matematicamente por:

$$\phi \leftarrow \phi - \alpha \frac{\delta L}{\delta \phi}. \quad (48)$$

Na prática, um modelo pode ter milhões de parâmetros, tornando a otimização computacionalmente custosa, uma das formas de diminuir o custo computacional é a utilização de diferentes algoritmos de otimização, como o *ADaptive Moment estimation* (ADAM), que combina as vantagens de duas extensões do gradiente descendente: o Momentum e o RMSProp. Calculando taxas de aprendizagem adaptativas individuais. Ele funciona através do cálculo das estimativas do primeiro momento (média dos gradientes) e do segundo momento (variância não centrada dos gradientes) para ajustar o passo de atualização de cada parâmetro.

Conforme apontado por Prince (2023) e Goodfellow, Bengio e Courville (2016), a escolha dos hiperparâmetros de treinamento de um modelo de rede neural é feita de forma empírica, através de treinamentos e testes do modelo de forma a verificar seu desempenho.

Figura 46 – Algoritmos de otimização de gradiente. a) Gradiente descendente ($\alpha = 0.05$) b) Gradiente descendente ($\alpha = 1$) c) Gradientes normalizados ($\alpha = 0.05$) d) ADAM ($\alpha = 0.05, \beta = 0.9, \gamma = 0.99$)



Fonte: Prince (2023).

2.7.3 Aprendizado Curricular

Em treinamento de redes neurais para a solução de problemas complexos, o algoritmo de otimização pode estagnar em mínimos locais sub-ótimos ou sofrer com uma convergência lenta. Para mitigar esse problema, Bengio *et al.* (2009) propuseram o conceito de *Curriculum Learning* (Aprendizado Curricular). Inspirada no processo cognitivo de aprendizado humano e no treinamento animal, essa estratégia propõe que o modelo não deve ser exposto imediatamente a todos os dados ou tarefas de forma aleatória. Em vez disso, o treinamento é organizado de forma que o sistema aprenda primeiro conceitos mais simples e, progressivamente, seja exposto a tarefas de maior complexidade.

É especialmente útil no aprendizado multitarefa, onde uma única rede deve aprender conceitos distintos. Ao iniciar o treinamento com tarefas mais fáceis, a função de custo apresenta uma superfície de erro mais comportada, permitindo que os parâmetros da rede se movam rapidamente para uma região promissora do espaço de busca antes de enfrentar a complexidade total do problema (GOODFELLOW; BENGIO; COURVILLE, 2016).

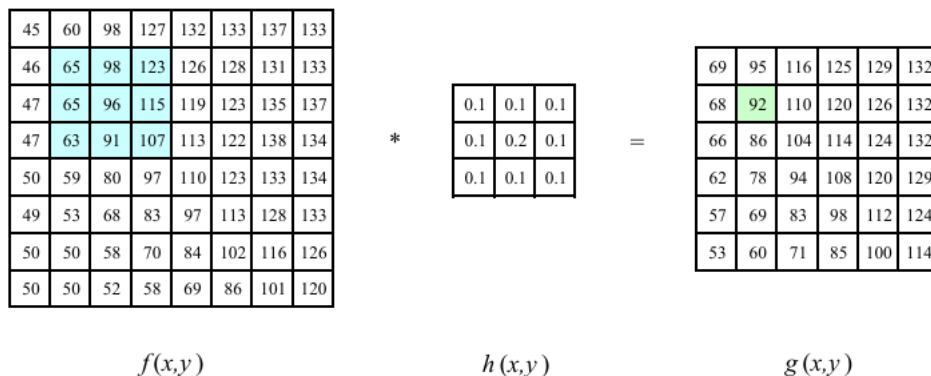
2.7.4 Redes Neurais Convolucionais

Uma Rede Neural Convolutiva (CNN) pode ser definida como "uma rede neural que usa convolução em vez de multiplicação de matrizes em pelo menos uma de suas camadas" (GOODFELLOW; BENGIO; COURVILLE, 2016, p 330). As CNNs são uma classe especializada de redes neurais, ideais para o processamento de dados com topologia em grade, como imagens. Sua principal diferença está na capacidade de explorar a correlação espacial e extrair características como bordas e texturas sem a necessidade de parametrização manual prévia.

Outra grande vantagem das CNNs sobre outros métodos, segundo Gu *et al.* (2017) é a sua capacidade de obter uma acurácia maior em grandes conjuntos de dados devido à sua capacidade de aplicar filtros que extraem as características das imagens e as combinam em um mapa de características que é aplicado, reduzindo o custo computacional.

A essência das CNNs está na aplicação de uma operação convolutiva (figura 47), uma operação que é feita passando um filtro (*kernel*) K de menor dimensão em uma imagem de entrada I . A convolução em uma (CNN) pode ser expressa pela equação 49, encontrada em Szeliski (2022).

Figura 47 – Operação de convolução do kernel $h(x,y)$ com a imagem $f(x,y)$, resultando em $g(x,y)$. Os pixels em azul representam os pixels que deram origem ao pixel em verde



Fonte: Szeliski (2022).

$$s(i,j,c_2) = \sum_{c_1 \in C_1} \sum_{(k,l) \in \mathcal{N}} w(k,l,c_1,c_2) x(i+k,j+l,c_1) + b(c_2), \quad (49)$$

onde:

\mathcal{N} representa o tamanho do kernel.

$s(i,j,c_2)$: É o valor resultante na posição (i,j) do mapa de características de saída c_2 .

$x(i+k,j+l,c_1)$: Representa os dados de entrada na posição $(i+k,j+l)$ no canal c_1 .

$w(k,l,c_1,c_2)$: São os pesos do filtro.

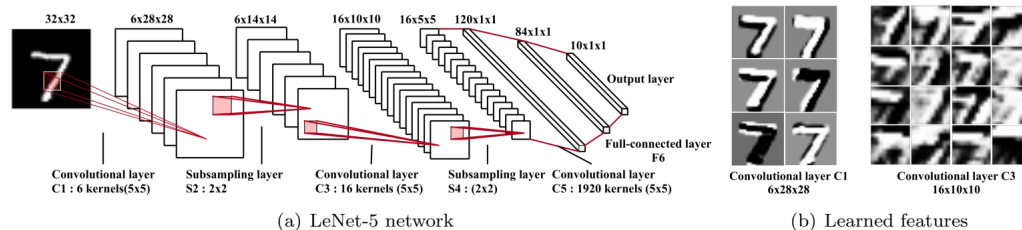
$b(c_2)$: É o *bias* do mapa de características c_2 .

Gu *et al.* (2017) aponta que, embora existam diversas variações de modelos de (CNNs), todas compartilham pelo menos 3 estruturas:

1. Camada Convolutiva: O que define uma CNN, nessa camada é feita a extração de características da imagem através da aplicação de um filtro de convolução. A complexidade computacional é reduzida devido ao compartilhamento de peso entre os neurônios.
2. Camada de *pooling*: Tem como objetivo reduzir gradualmente o mapa de características e da imagem, reduzindo assim o custo computacional. É colocada entre duas camadas convolucionais.
3. Camada totalmente conectada: Todos os neurônios dessa camada são conectados a todos os neurônios da camada anterior, é usada para raciocínios de alto nível que gerem informações globais sobre as características da imagem

A Figura 48 mostra uma arquitetura CNN típica, a LeNet-5, bastante utilizada em tarefas de reconhecimento de caracteres, composta por camadas convolucionais, de *pooling* e densas em sequência.

Figura 48 – a) Arquitetura de uma CNN LeNet-5, utilizada em tarefas de reconhecimento de dígitos. b) Visualização das features aprendidas pela rede.



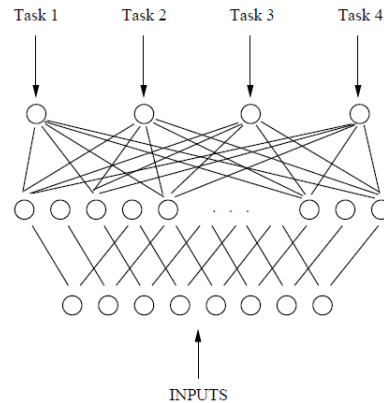
Fonte: Gu *et al.* (2017).

2.7.5 Aprendizado Multitarefa

Geralmente os modelos de redes neurais são projetados para resolver uma única tarefa isolada (*Single Task Learning* (STL)). Em contraste, o Aprendizado Multitarefa (*Multi Task Learning* (MTL)) é um paradigma de transferência indutiva que propõe o treinamento simultâneo de um modelo em múltiplas tarefas relacionadas (CARUANA, 1997).

A premissa fundamental do MTL está no compartilhamento de parâmetros. A figura 49 ilustra essa topologia de rede neural, composta por camadas ocultas compartilhadas (*backbone*) que alimentam camadas específicas para cada tarefa (*heads*). É uma abordagem que reduz o risco de *overfitting*, pois o número de parâmetros disponíveis para se ajustar a cada tarefa individual é reduzido, enquanto a quantidade de dados de treinamento efetivos aumenta para as camadas compartilhadas.

Figura 49 – Esquemático de uma rede multitarefa, o *backbone* alimenta sub-redes específicas para cada tarefa



Fonte: Caruana (1997).

2.7.5.1 Métricas de avaliação

A avaliação do desempenho de uma rede neural depende da tarefa a qual ele é submetido, para tarefas de classificação binária, essa análise se baseia na matriz de confusão.

Ela correlaciona as classes verdadeiras dos eventos simulados com as classes previstas pelo modelo, como na figura 50:

Figura 50 – Matriz de confusão

		classe predita	
		P	N
classe real	P	verdadeiro positivo (TP)	falso negativo (FN)
	N	falso positivo (FP)	verdadeiro negativo (TN)

Fonte: Adaptado de Raschka e Mirjalili (2017).

Onde temos:

Verdadeiro Positivo (**T**True **P**Positive): Eventos corretamente classificados.

Verdadeiro Negativo (**T**True **N**egative): Eventos corretamente rejeitados.

Falso Positivo (**F**alse **P**ositive): Eventos incorretamente classificados.

Falso Negativo (**F**alse **N**egative): Eventos incorretamente rejeitados

A partir da matriz de confusão, é possível definir a Acurácia (ACC) do modelo através da equação:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}. \quad (50)$$

Além de outras métricas, como a taxa de verdadeiros positivos (TPR) e falsos positivos (FPR), uma métrica utilizada para classes desbalanceadas:

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} \quad (51)$$

$$TPR = \frac{TP}{P} = \frac{TP}{FN + TP}$$

A taxa de verdadeiros positivos também é chamada de *Recall*, outra métrica é a precisão (PRE):

$$PRE = \frac{TP}{TP + FP}. \quad (52)$$

Na prática costuma-se utilizar uma combinação entre Recall e Precisão, chamada de F1-Score:

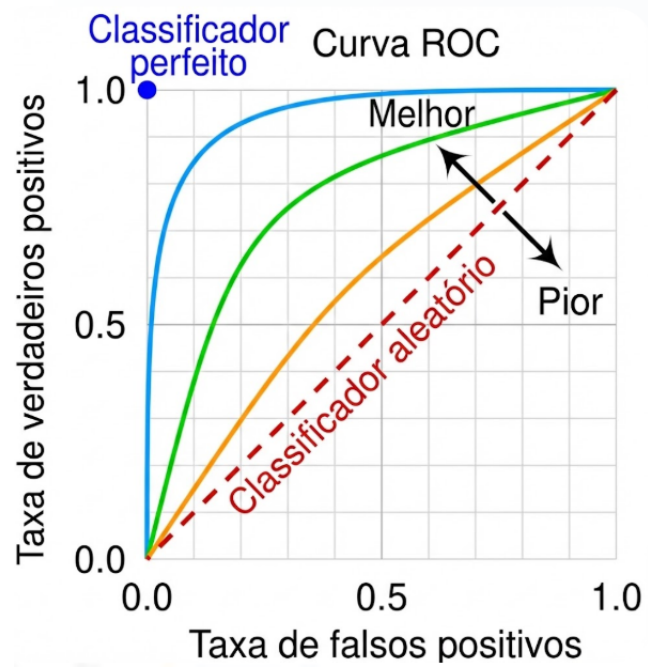
$$F1 = 2 \frac{PRE \times REC}{PRE + REC}. \quad (53)$$

Em algumas aplicações é necessário o uso de métricas complementares. A saída da rede neural é uma probabilidade contínua de que o evento faça parte de uma das classes. Ao variar o limiar de corte desse score, altera-se a taxa de verdadeiros e falsos positivos.

A Curva característica de operação do receptor (*Receiver Operating Characteristic* (ROC)) é o gráfico da taxa de verdadeiros positivos versus a taxa de falsos positivos para todos os limiares possíveis. A Área Sob a Curva (*Area Under the Curve* (AUC)) resume o desempenho do classificador em um único escalar entre 0,5 (classificação aleatória) e 1 (classificação perfeita).

Diferentemente de distribuições puramente Gaussianas, os erros de reconstrução em IACTs apresentam caudas longas e assimetrias devido a flutuações estocásticas do chuveiro e imagens truncadas nas bordas da câmera. Por esse motivo, para as tarefas de regressão usa-se como métrica padrão o intervalo de contenção de 68%. Para a energia, define-se como a metade da largura do intervalo centrado em 0 que contém 68% dos eventos. Para a direção, utiliza-se o raio angular θ_{68} que contém 68% dos eventos reconstruídos em torno da posição verdadeira da fonte.

Figura 51 – Curva ROC. A diagonal mostra o desempenho de um classificador aleatório. Três exemplos de classificadores (azul, laranja e verde) são apresentados.



Fonte: Adaptado de Lee (2018).

3 MÉTODOS

3.1 Simulação de eventos

O desenvolvimento de modelos de *Deep Learning* para a astrofísica de partículas depende da qualidade e fidelidade dos dados utilizados no treinamento. Dado que fontes reais de raios gama não fornecem o rótulo verdadeiro do tipo, energia ou direção de cada partícula, a utilização de simulações de Monte Carlo torna-se uma ferramenta obrigatória, reforçado pelo fato de que o CTAO ainda não está 100% operacional.

Esta seção descreve o *pipeline* de produção de dados, que consiste na simulação da física das cascatas atmosféricas iniciadas por raios gama (sinal) e prótons (ruído), utilizando o CORSIKA, seguida pela simulação da resposta instrumental do telescópio com o `sim_telarray`.

3.1.1 CORSIKA

A geração de EASs simulados foi realizada utilizando o CORSIKA 77500, compilado via CORSIKA CONFIGURATION UTILITY - COCONUT, com um conjunto específico de opções de compilação selecionadas de acordo com as necessidades do nosso projeto, que podem ser consultadas nos documentos CORSIKA Guide e CORSIKA Physics, disponibilizados por Heck *et al.* (1998).

3.1.2 Compilação

Interações de altas energias: Para interações com energia primária acima de 80 GeV, utilizou-se o modelo *Quark Gluon String model with JETs* (versão II-04), que incorpora correções de interação baseadas na troca de Pomerons e foi calibrado com dados recentes do LHC (*Large Hadron Collider*).

Interações de baixas energias: Para energias abaixo de 80 GeV, foi selecionado o modelo *Ultra-relativistic Quantum Molecular Dynamics (URQMD 1.3cr)*, que realiza uma simulação detalhada das colisões hádron-núcleo, essencial para a estimativa da componente muônica e hadrônica de baixa energia.

Geometria do detector: A opção "VOLUMEDET" foi selecionada, ela adequa a simulação à geometria esférica dos IACTs.

Distribuição Longitudinal da Luz Cherenkov: Foi selecionada a opção "*Photons counted only in the step where emitted*", assim o programa contabiliza os fótons no passo de simulação onde foram gerados, em vez de propagá-los até o nível de observação, diminuindo o custo computacional que teria para fazer o rastreamento.

Independência do Comprimento de Onda no Ângulo de Emissão: A opção selecionada foi a "*Emission angle is wavelength independent*". Estima-se que a variação na abertura do cone seja inferior a $0,03^\circ$, portanto menor que a resolução angular de um pixel ($\sim 0,1^\circ$). Compilando com essa opção reduzimos a complexidade do cálculo de Monte Carlo.

IACTEXT: Foi selecionada a opção de não armazenar as partículas no nível do detector (*Particle at detector level not stored*) no arquivo de saída. Assim o CORSIKA descarta as informações das partículas secundárias que atingem a superfície, registrando no arquivo de saída apenas os fótons Cherenkov que chegaram à região de detecção.

Também foram compiladas opções adicionais do CORSIKA:

CERENKOV e IACT: A opção CERENKOV habilita a produção de fótons Cherenkov na atmosfera. A opção IACT, especificamente, modifica a saída padrão do CORSIKA, agrupando os fótons em pacotes e registrando-os em um formato de arquivo eventio de Bernlohr.

ATMEXT (*Atmosphere Extension*): Permite a substituição da atmosfera padrão do CORSIKA por perfis atmosféricos tabulados externos, o que possibilita o uso de dados reais de densidade e índice de refração atmosférica do sítio de La Palma para uma simulação mais fiel da interação com a atmosfera.

VIEWCONE: Para a geração de eventos cuja direção de chegada é sorteada aleatoriamente dentro de um cone em torno do eixo óptico. Essencial para a geração de eventos em todas as posições possíveis do campo de visão da câmera.

3.1.3 Arquivos de *Input*

Para realizar as simulações, é preciso criar um arquivo de entrada onde são definidos os parâmetros da simulação para cobrir diferentes condições observacionais e garantir uma variação estatística para a geração do *dataset* para o treinamento da rede neural. É nesse arquivo onde são definidos parâmetros como tipo da partícula primária (1 para fóton e 14 para próton), posicionamentos dos detectores para simular a distribuição geográfica dos 4 LSTs, formatos de saída e outros detalhes pertinentes à simulação, as principais escolhas para a geração do dataset estão detalhadas abaixo:

Intervalo de energia: O intervalo de energia simulado foi entre 5 GeV e 300 GeV. Um intervalo um pouco maior do que a sensibilidade nominal dos LSTs, assim a rede consegue modelar efeitos de limiar.

Geometria de Observação e Atmosfera: As simulações foram realizadas em três configurações de ângulo zenital (θ): 20° , 40° e 60° , com um modelo atmosférico obtido para o sítio do CTAO-Norte (La Palma) com a refração atmosférica ativada.

Reutilização de Eventos e Área de Impacto (CSCAT): Visando a eficiência computacional foi utilizada a técnica de *Core Scattering* (CSCAT), que reutiliza o desenvolvimento de um chuveiro, mas posicionando-o em locais diferentes. Para raios gama cada chuveiro foi reutilizado 10 vezes em observações pontuais e 20 vezes em observações difusas, com impactos

distribuídos em um raio de 500 m, já para raios cósmicos, cada chuva foi reutilizado 40 vezes, com impactos distribuídos em um raio de 800 m.

Geometria da Fonte: Foi utilizada a opção VIEWCONE para simular uma distribuição isotrópica ao redor da fonte, com uma abertura de 10° .

3.1.4 O pacote sim_telarray

A segunda etapa da criação do banco de dados consiste no processamento dos arquivos de saída do CORSIKA através do pacote sim_telarray, que faz o rastreamento detalhado dos fótons Cherenkov (*ray-tracing*) através do telescópio e simula a resposta completa da aquisição eletrônica.

Para reproduzir as condições de observação dos telescópios LSTs no sítio norte (La Palma), utilizou-se o arquivo de configuração padrão da colaboração usado na geração de dados de testes e validação mais recente, denominada Prod6 (OBSERVATORY, 2024). Esse arquivo define todo o instrumento que fará a observação: a geometria do refletor parabólico segmentado, incluindo as curvas de refletividade e as imperfeições de alinhamento, degradação da superfície refletora, sombreamento causado pelo suporte, geometria dos pixels da câmera, eficiência quântica dos fotomultiplicadores, eficiência dos tubos de Winston, ruído eletrônico, etc.

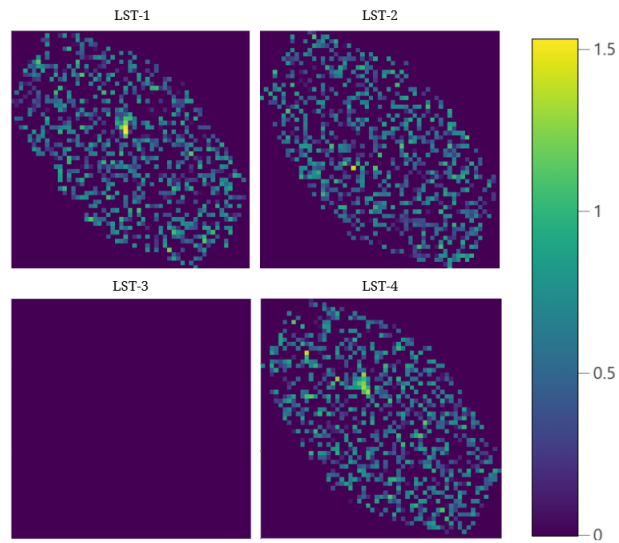
O resultado das simulações é um arquivo comprimido no formato EVENTIO, um padrão adotado pela colaboração CTAO que permite o armazenamento eficiente de grandes volumes de dados. Organizados internamente em blocos contendo informações detalhadas da simulação, desde a imagem gerada até os parâmetros físicos que a geraram.

3.2 Conjunto de dados

O conjunto de dados (*dataset*) compreende eventos compostos por imagens de chuvis iniciados por raios gama (Figura 52) e prótons (Figura 53) na configuração de quatro LSTs localizados no CTAO norte, minimamente processadas, bem como parâmetros físicos do EAS necessários para o treinamento da rede.

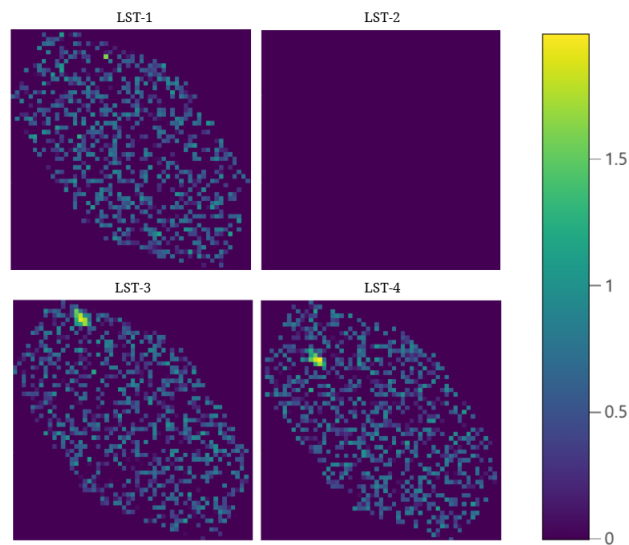
Ao todo, foram utilizados 500 mil eventos, divididos em 400 mil eventos para o treinamento e 100 mil eventos para a validação da rede neural. Como pode ser observado no perfil da distribuição energética dos eventos utilizados para o treinamento da rede neural (Figura 54) e na validação (Figura 55), o conjunto de dados apresenta um desbalanceamento nas classes, sendo dominado por raios gama pontuais (representando cerca de 72% dos dados). Em contrapartida, os prótons e os raios gama difusos constituem as classes minoritárias ($\sim 16\%$ e $\sim 12\%$, respectivamente). Vale ressaltar que esse desbalanceamento é uma implicação direta da Física do desenvolvimento do EAS e da sensibilidade do instrumento. Chuvis hadrônicos iniciados

Figura 52 – Imagens de um EAS iniciado por gama (~ 60 GeV) a 40°



Fonte: Autoria própria.

Figura 53 – Imagens de um EAS iniciado por próton (~ 60 GeV) a 40°

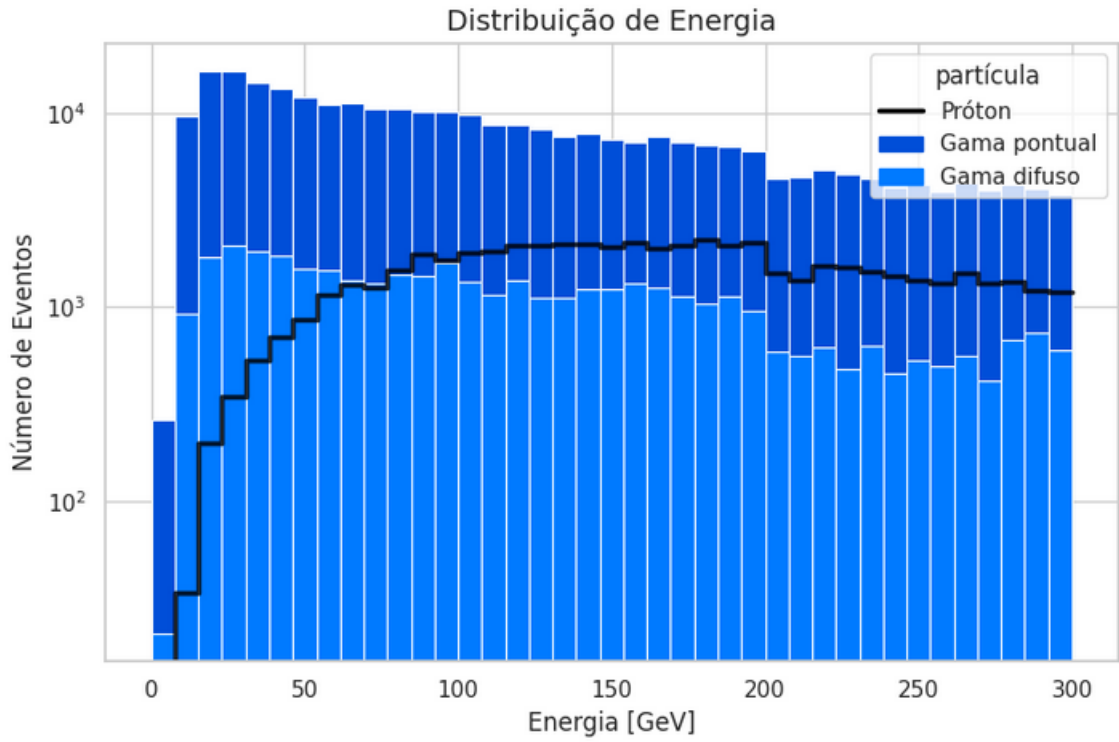


Fonte: Autoria própria.

por prótons são mais dispersos e por vezes não superam os limiares de luz Cherenkov necessários para acionar a câmera. De maneira análoga, raios gama difusos sofrem com efeitos de bordas, reduzindo sua taxa de acionamento em comparação aos gamas pontuais alinhados ao eixo óptico.

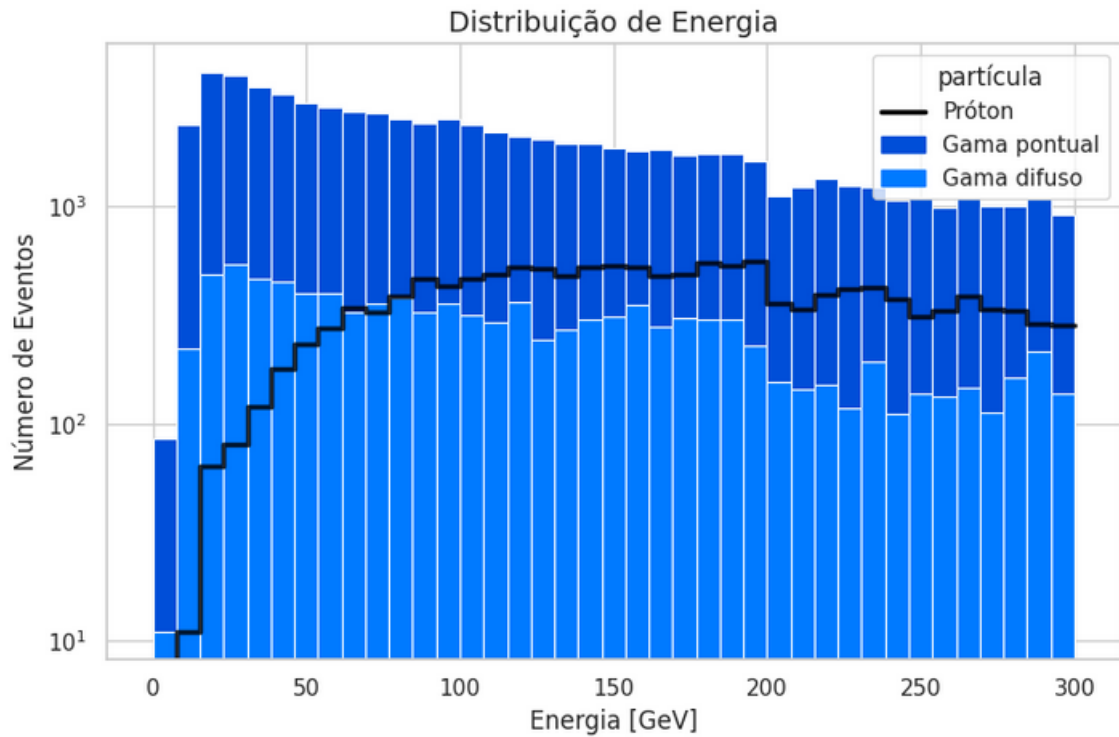
Para contornar esse efeito, a métrica de erro do modelo foi adaptada utilizando a função Focal Loss, conforme apresentado na Seção 2.7.2.0.1, obrigando a rede a concentrar a atualização de seus pesos nas amostras minoritárias e de difícil classificação geométrica.

Figura 54 – Distribuição energética dos eventos de treino



Fonte: Autoria própria.

Figura 55 – Distribuição energética dos eventos de validação



Fonte: Autoria própria.

3.2.1 Obtenção dos dados e Pré-Processamento

A etapa seguinte consistiu na leitura dos dados brutos gerados pelo `sim_telarray`, utilizando as bibliotecas `pyeventio` (CTA Observatory, 2024c) e `DL1-Data-Handler` (CTA Observatory, 2024b) para efetuar a extração e o mínimo pré-processamento dos dados necessários para o treinamento da rede neural.

Foram selecionados dados relevantes para a reconstrução do evento, tanto informações da simulação, que serão utilizadas como rótulos verdadeiros para o treinamento, quanto as informações de aquisição do sinal pelo telescópio: Sinal gerado pela câmera, rótulos de energia e direção da partícula primária, direção de apontamento do telescópio, ponto de impacto do chuveiro no chão, tipo da partícula, altura da primeira interação e número de identificação do evento.

Em vez de utilizar a forma de onda completa digitalizada, vinda das amostras FADC, foi utilizado o sinal da integração de carga do canal de alto ganho dos PMTs em uma janela temporal fixa de 8 ms ao redor do pico do sinal, essa escolha maximiza a Relação Sinal-Ruído, pois integra apenas a região onde o pulso de luz Cherenkov está presente, rejeitando o ruído eletrônico e de fundo de céu do restante do período de aquisição.

O nível de base instrumental (pedestal), que representa o deslocamento DC da eletrônica, foi subtraído antes da integração, garantindo que o sinal represente apenas a carga induzida pelos fótons Cherenkov incidentes.

O valor integrado em contagens digitais (ADC Counts) foi convertido para o número de fotoelétrons (p.e.) através da aplicação de fatores de calibração absoluta, visando normalizar a resposta dos diferentes pixels e telescópios, utilizando um fator de conversão de 40 ADC/p.e..

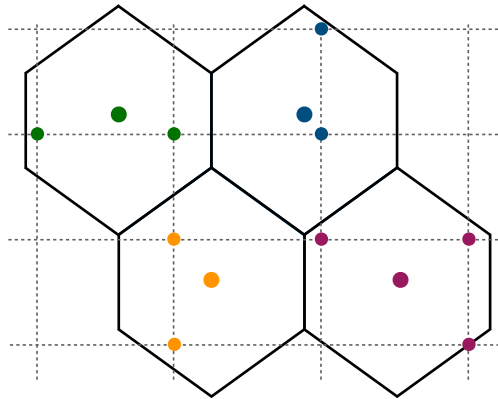
Após a calibração física, as imagens passaram por transformações matemáticas para adequá-las à topologia e à dinâmica de aprendizado das CNNs, primeiro aplicando um corte nos valores negativos resultantes da subtração do pedestal em pixels com apenas flutuação de ruído.

A faixa de fotoelétrons em uma detecção de evento Cherenkov atmosférico pode variar de poucos fotoelétrons na borda do chuveiro e no ruído de céu, até dezenas de milhares de fotoelétrons no núcleo. Essa grande amplitude poderia causar uma explosão de gradiente durante o processo de aprendizado. Para lidar com isso, aplicou-se uma transformação logarítmica na intensidade dos pixels, dessa forma a rede neural é capaz de extrair características tanto das bordas quanto do núcleo do chuveiro, sem que este último seja dominante na função de perda.

Outro ponto de atenção é que CNNs tipicamente operam com matrizes quadradas e a imagem obtida pelo LST possui pixels hexagonais, embora existam estudos sobre o uso de um kernel hexagonal (NIETO *et al.*, 2019; ZHAO *et al.*, 2021), optamos pela utilização do padrão bem estabelecido, por conta disso foi realizada uma transformação geométrica nas imagens, utilizando uma técnica de reamostragem para projetar a imagem hexagonal da câmera em uma malha cartesiana de 55×55 pixels.

Esta tarefa foi executada utilizando a técnica de interpolação por Vizinheiro Mais Próximo, ilustrada na figura 62 onde cada pixel assume o valor do hexágono com o centro mais próximo, preservando a natureza discreta e a estatística de Poisson dos contadores de fótons e mantendo a integridade das bordas.

Figura 56 – Representação do método de mapeamento por vizinho mais próximo



Fonte: Nieto et al. (2019).

As simulações de EAS geram um grande volume de dados brutos, ultrapassando facilmente a ordem de Terabytes. Para mitigar esse gargalo computacional foi utilizado o formato hierárquico HDF5 para o armazenamento da base de dados. Esse formato suporta nativamente o armazenamento de tensores, permite acesso randômico e leitura paralela dos dados, além de possuir algoritmos internos de compressão.

A base de dados contém os dados relevantes para o treinamento, englobando os dados de entrada, os rótulos *ground truth* e dados auxiliares:

Dados de entrada:

1. Imagens estereoscópicas: As imagens de cada evento são alocadas em um tensor com formato $(n_{eventos}, 4, 55, 55)$, cada tensor contém um evento distinto e as informações de observação pelos telescópios. Caso um telescópio não tenha observado o EAS, a técnica *zero-padding* é empregada e sua imagem consistirá em uma matriz de zeros.
2. Apontamento do telescópio: Contém as coordenadas do ponto no céu para onde o telescópio está apontado (tel_alt, tel_az) . Garantindo uma reconstrução mais fiel da energia do chuveiro e de sua origem.
3. Máscara de Gatilho: Um vetor binário indicando quais telescópios efetivamente participaram da detecção.

Rótulos *ground truth*:

1. Tipo da partícula: Para efeitos de classificação, armazena qual partícula primária iniciou o chuveiro.

2. Energia: Energia real da partícula primária.
3. Direção: Coordenadas (*altitude*, *azimute*) de chegada da partícula, no caso dos raios gama, também indica a direção da fonte.

Dados auxiliares:

1. Altura da primeira interação: Local que marca o início do desenvolvimento do EAS na atmosfera.
2. Ponto de impacto: Coordenadas ($Core_x$, $Core_y$) do ponto de impacto do eixo do EAS na superfície.

3.3 Arquitetura da Rede Neural

Para processar as imagens estereoscópicas dos quatro telescópios LST e realizar simultaneamente a classificação e a reconstrução da energia e direção da partícula primária, foi desenvolvida uma arquitetura de topologia siamesa baseada em blocos convolucionais residuais equipados com módulos de *Squeeze-and-Excitation*. Com a introdução de mecanismos de atenção e fusão estereoscópica, além da adoção de uma estratégia de Pooling híbrido guiada por princípios físicos.

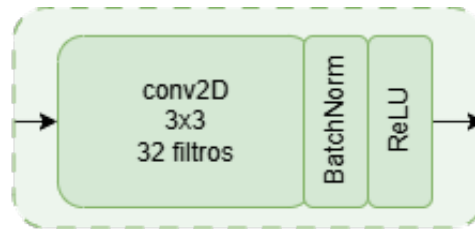
3.3.1 Backbone siamês

A entrada do sistema consiste nas imagens dos quatro telescópios LST. Para garantir a invariância translacional e a eficiência paramétrica, optou-se por uma arquitetura com um *backbone* siamês, composta por quatro instâncias idênticas de uma rede convolucional profunda que compartilham o mesmo conjunto de pesos. O processamento inicial de cada imagem é realizado pelo estágio da Figura 57, conhecido como *Stem*, seguido por três blocos convolucionais residuais (Figura 58) progressivos com 64, 128 e 256 filtros. Cada bloco residual possui duas convoluções sucessivas de 3 x 3 que permitem a extração de padrões morfológicos complexos, enquanto reduz a suscetibilidade ao *overfitting* com uma maior eficiência computacional se comparado ao uso de *kernels* maiores.

Cada bloco residual possui integrado o mecanismo de *Squeeze-and-Excitation* (SE), representado na Figura 59, o SE permite que a rede aprenda a importância relativa de cada mapa de características, enfatizando componentes morfológicos relevantes do chuveiro atmosférico e anulando ruídos de fundo através de uma operação de excitação via ativação sigmoide. O uso da conexão de atalho facilita o fluxo de gradientes e a preservação da identidade da imagem, permitindo que o modelo aprenda filtros morfológicos complexos.

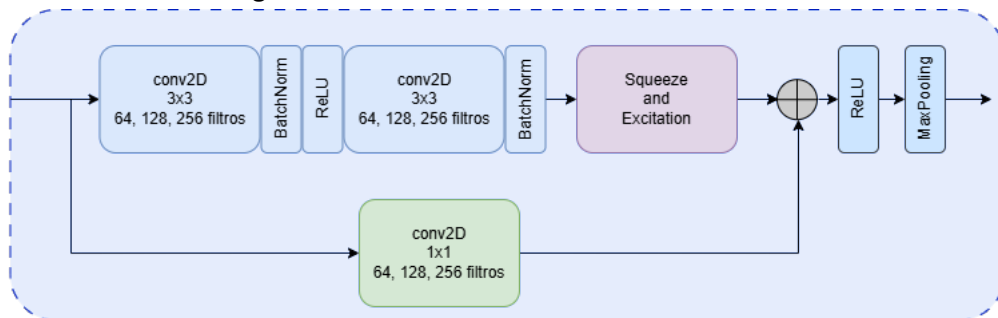
Foi implementada uma camada de *Multi-Head Self-Attention*, ilustrada na Figura 60, em cada ramo siamês após os blocos convolucionais, ela permite que a rede consiga capturar de-

Figura 57 – Bloco STEM, responsável por extrair as primeiras características e preparar a imagem para a rede



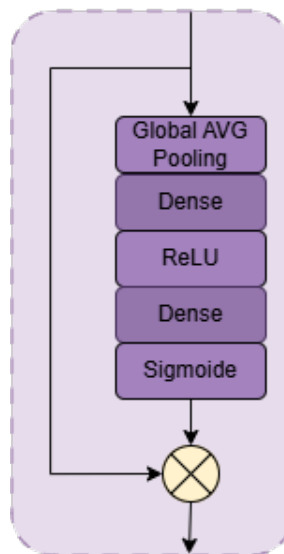
Fonte: Autoria própria.

Figura 58 – Blocos convolucionais residuais



Fonte: Autoria própria.

Figura 59 – Bloco de Squeeze-and-Excitation

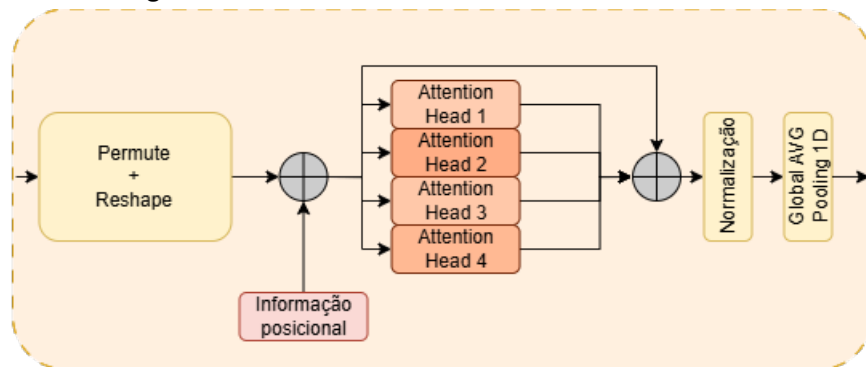


Fonte: Autoria própria.

pendências globais na imagem. Para isso, os mapas de características finais são linearizados em uma sequência de vetores de características e acrescidos de uma representação posicional aprendível, que preserva a informação geométrica da grade original. O mecanismo de atenção é configurado com quatro cabeças de atenção, que permitem que o modelo correlacione diferentes regiões da imagem do chuva.

Após o mecanismo de atenção, os vetores de características resultantes do processamento de cada telescópio são integrados via fusão por concatenação. Neste estágio, são injeta-

Figura 60 – Mecanismo de *multi-head self-attention*



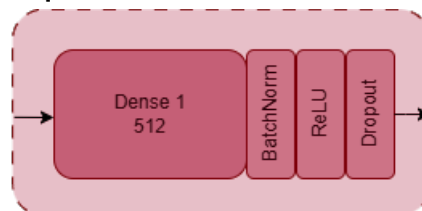
Fonte: Autoria própria.

das as variáveis auxiliares de apontamento do telescópio (altitude e azimute). A inserção desses dados diretamente antes das camadas densas permite que a rede correlacione a morfologia da imagem com a orientação geométrica do sistema, facilitando o aprendizado do processo de reconstrução estereoscópica do chuvaire.

3.3.2 Gargalo (*Neck*)

O vetor fundido alimenta um tronco, composto por três camadas densas (Figura 61) com 512, 256 e 128 unidades, respectivamente. Este módulo atua como um gargalo, sintetizando as informações visuais e de apontamento antes de distribuí-las para as quatro *heads*.

Figura 61 – Esquemático do bloco denso utilizado no *neck*



Fonte: Autoria própria.

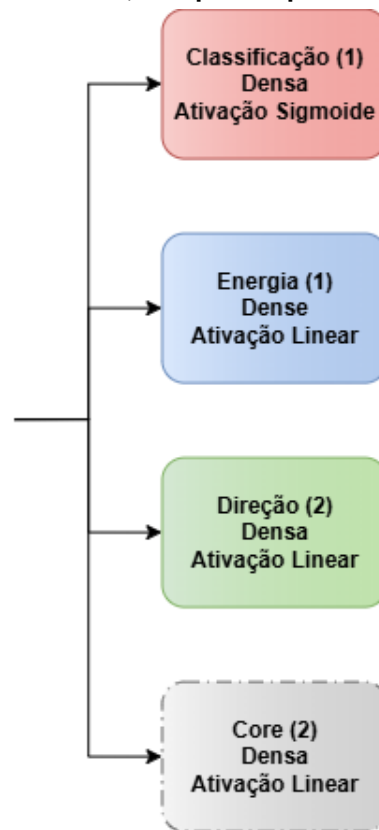
3.3.3 Decodificadores (*Heads*)

A partir do vetor global de características fundidas, a rede bifurca-se em quatro ramificações independentes (*Heads*), cada uma especializada em uma tarefa física distinta, configurando a estratégia de Multitask Learning:

1. Classificação Gama-Hádron: Composta por uma camada densa com função de ativação Sigmoide. A saída é um escalar que representa a probabilidade de o evento ser um raio gama.

2. Reconstrução da Energia: Composta por uma camada densa com ativação linear na saída. O objetivo é a regressão do logaritmo da energia primária.
3. Reconstrução da Origem: Composta por uma camada densa projetada para fazer a regressão das coordenadas de origem da fonte no céu.
4. Localização do impacto: Utiliza uma camada densa para estimar as coordenadas (X, Y) do centro do chuveiro no solo. Embora secundária, a inclusão desta tarefa força o backbone a aprender a paralaxe estéreo e a relacionar a energia do evento e a sua distância do telescópio.

Figura 62 – As 4 heads, compostas por camadas densas



Fonte: Autoria própria.

3.3.4 Pooling Híbrido

A estratégia adotada para a redução da dimensionalidade foi o uso do pooling híbrido, sendo utilizado o Average Pooling para as tarefas de classificação e reconstrução da direção e o Sum Pooling posteriormente para a reconstrução da energia.

3.3.4.1 Global Average Pooling

Calcula a média dos mapas de ativação. Este método preserva a informação espacial e morfológica, sendo ideal para as tarefas de classificação (Gama/Hádron) e reconstrução geométrica (Direção), onde a intensidade absoluta é menos relevante que a topologia do evento

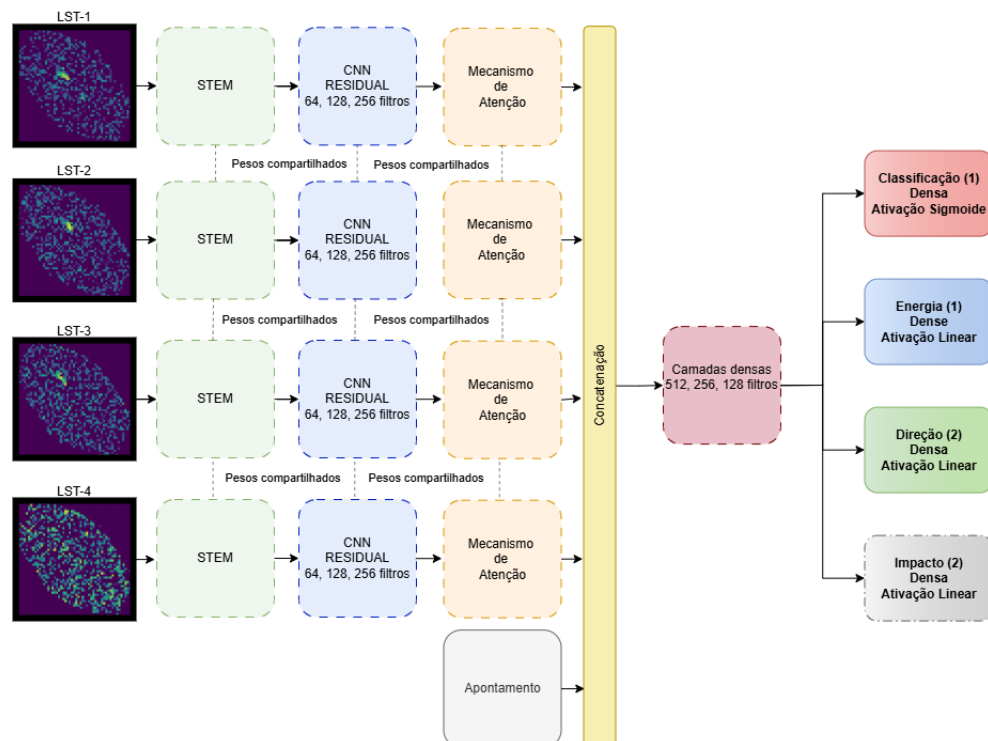
3.3.4.2 Global Sum Pooling

Calcula a soma total das ativações. Esta escolha é motivada pelo princípio da calorimetria: a energia da partícula primária é proporcional ao número total de fótons Cherenkov produzidos na atmosfera. Ao usar a soma, a rede preserva a fotometria.

3.4 Detalhes de Implementação

A arquitetura proposta, representada integralmente de forma simplificada na Figura 63, foi implementada utilizando o *framework* TensorFlow 2.14.0 com a API Keras. O modelo final apresentado aqui é fruto de um refinamento contínuo executado em etapas.

Figura 63 – Representação simplificada da rede neural



Fonte: Autoria própria.

Nas primeiras etapas do treinamento do modelo foi utilizada a ponderação automática da função de perda baseada na incerteza homocedástica de Kendall, Gal e Cipolla (2017). Isso permitiu que fosse alcançada uma região próxima ao mínimo global levando em consideração

todas as tarefas. Após essa etapa, foi realizado um *fine tuning* da tarefa de energia. Essa forma de calcular a perda falhava para a reconstrução da direção da partícula, com a rede aumentando artificialmente o parâmetro de incerteza para minimizar a perda global, priorizando tarefas mais fáceis.

Para corrigir esse problema, foi realizada uma alteração na função de perda, passando a ser uma ponderação estática, na qual pesos manuais foram atribuídos à função de perda, penalizando severamente os erros de direção e ponto de impacto com fatores de elevados. Forçando o *backbone* a aprender representações geométricas e correlacionar as 4 entradas de forma estereoscópica.

O ajuste fino da reconstrução da energia apresenta uma particularidade, sendo necessário fazer a soma dos pixels no *pooling* para efeitos de totalização dos fotoelétrons. Foi feito o transplante dos pesos do *backbone* treinado geometricamente, retreinando apenas as novas camadas densas.

Após a estabilização das tarefas geométricas e energéticas, foi realizado um ajuste fino de todas as camadas, permitindo que todos os pesos da rede fossem ajustados com uma taxa de aprendizado reduzida (10^{-6}). Equilibrando a precisão angular e a calibração energética e culminando no modelo apresentado aqui.

3.5 Métricas de Avaliação

Para quantificar o desempenho do modelo proposto, foram adotadas métricas padrão para análise de IACTs:

3.5.1 Métricas de Classificação (Separação Gama/Hádron)

A capacidade do modelo em rejeitar o fundo de raios cósmicos (hádrons) e preservar os eventos de raios gama (sinal) foi avaliada utilizando métricas clássicas para problemas de classificação binária:

3.5.1.1 Curva ROC e AUC

A curva ROC ilustra o compromisso entre a Eficiência de Sinal (Taxa de Verdadeiros Positivos) e a Eficiência de Fundo (Taxa de Falsos Positivos). A Área Sob a Curva (AUC) fornece um índice global de separabilidade onde 1.0 indica discriminação perfeita e 0.5 equivale a uma escolha aleatória.

3.5.1.2 Matriz de Confusão

Uma representação tabular que explicita a distribuição das predições em quatro categorias: Verdadeiros Positivos (Gamas corretamente identificados), Verdadeiros Negativos (Prótons corretamente rejeitados), Falsos Positivos (Prótons contaminantes) e Falsos Negativos (Gamas perdidos). Esta ferramenta é essencial para visualizar o viés do classificador.

3.5.1.3 Precisão, Recall e F1-Score

Dada a importância crítica de obter uma amostra pura de fótons para a análise espectral, avaliam-se também a Precisão, que quantifica a pureza da amostra classificada como gama (baixa contaminação hadrônica) e o Recall (ou Eficiência), que quantifica a fração de fótons reais recuperados pelo detector. O F1-Score é a média harmônica entre ambos, oferecendo uma métrica única de equilíbrio.

3.5.1.4 Acurácia

Define a fração global de acertos (para gamas e hádrons) considerando um limiar de decisão fixo de 0.5. Embora reportada para fins de comparação com a literatura de aprendizado de máquina, seu valor precisa ser interpretado com cautela devido à natureza desbalanceada dos eventos.

3.5.2 Métricas de Reconstrução de Energia

A performance da estimativa de energia é avaliada exclusivamente sobre o subconjunto de eventos de raios gama. O erro relativo de energia para cada evento é definido como:

$$\epsilon = \frac{E_{pred} - E_{true}}{E_{true}}. \quad (54)$$

Diferente de abordagens estatísticas convencionais que assumem uma distribuição normal, as distribuições de erro em IACTs frequentemente apresentam caudas não-gaussianas. Portanto, adotam-se métricas distintas.

3.5.2.1 Viés de Energia

Definido como a média da distribuição do erro relativo ϵ em cada nível de energia. Indica a tendência sistemática do modelo em subestimar ou superestimar a energia da partícula primária.

3.5.2.2 Resolução Energética

Definida como a metade da largura do intervalo centrado em 0 que contém 68% da distribuição dos erros relativos. Esta métrica, conhecida como região de 68% de contenção, é o padrão da área para mitigar o impacto de *outliers* na avaliação de desempenho.

$$Res_E = \frac{Q_{84}(\epsilon) - Q_{16}(\epsilon)}{2}. \quad (55)$$

Onde Q_{84} e Q_{16} representam, respectivamente, o 84° e o 16° percentil da distribuição de erros.

3.5.3 Matriz de Migração

Para visualizar a correlação e a dispersão entre a energia estimada e a energia real, utiliza-se a Matriz de Migração. Esta ferramenta consiste em um histograma bidimensional normalizado que plota a energia real no eixo das abscissas pela energia predita no eixo das ordenadas.

Uma reconstrução ideal resulta em uma linha diagonal perfeita ($y=x$). O alargamento da distribuição em torno da diagonal ilustra visualmente a resolução energética, enquanto desvios sistemáticos da diagonal indicam o viés de reconstrução do modelo em determinadas faixas de energia.

3.5.4 Métricas de Reconstrução da Direção

A precisão na determinação da origem do raio gama no céu é avaliada através da separação angular (θ) entre a direção reconstruída e a direção verdadeira. A Resolução Angular (θ_{68}) é definida como o raio de contenção de 68% da Função de Espalhamento Pontual (PSF). É o valor angular ψ tal que 68% dos eventos reconstruídos possuem erro $\theta < \Psi$. Esta é a métrica padrão para IACTs e permite quantificar a capacidade do instrumento em resolver a morfologia de fontes extensas e reduzir a contaminação de fundo em fontes pontuais.

4 RESULTADOS E DISCUSSÃO

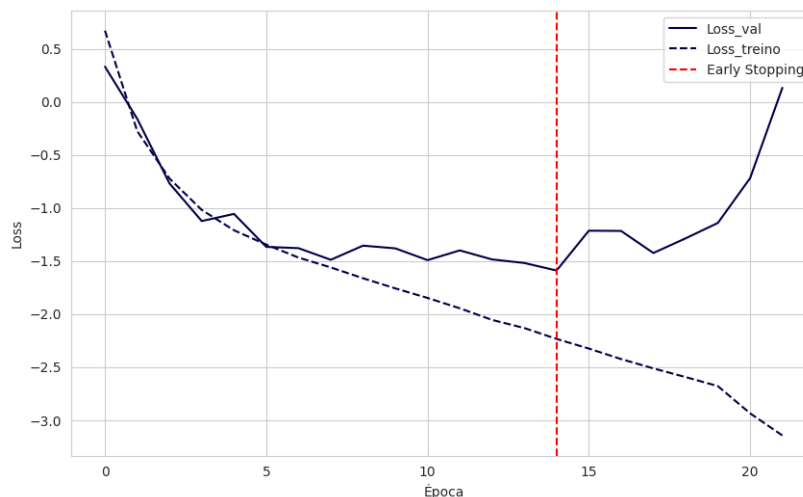
Neste capítulo, serão apresentados e discutidos os resultados obtidos pela rede neural proposta nas tarefas de classificação e reconstrução de parâmetros de EAS. A análise divide-se em quatro eixos fundamentais: a evolução e convergência do treinamento, a eficácia da separação entre sinais de raios gama e o ruído hadrônico, a precisão calorimétrica na estimativa de energia e, por fim, a resolução angular sob diferentes configurações geométricas.

4.1 Análise da Evolução do Treinamento e Convergência do modelo

A convergência de uma arquitetura multitarefa, é um desafio de otimização complexo. A estratégia de treinamento em múltiplos estágios revelou-se de grande importância para a estabilidade da rede neural, sem isso, corre-se o risco do modelo estagnar em mínimos locais relacionados às tarefas mais fáceis

A Figura 64 detalha o comportamento da função de perda global durante o primeiro estágio de treinamento, onde foi aplicada a ponderação automática via Incerteza Homocedástica, usado como critério de parada para o *early stopping* para prevenir o *overfitting*. É possível observar que a rede aprende a classificar os eventos em poucas épocas e começa a decorá-los.

Figura 64 – Valores da Loss Global (Incerteza Homocedástica) para o treino e validação no primeiro estágio de treinamento

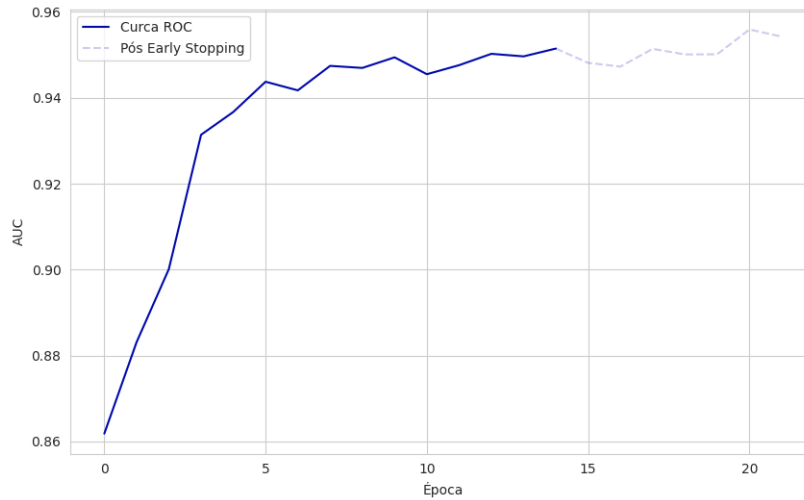


Fonte: Autoria própria.

A classificação binária se apresenta como a tarefa mais fácil, alcançando valores significativos já nas primeiras épocas, como demonstra a Figura 65.

Enquanto energia e direção são tarefas mais complexas, necessitando de mais tempo de treinamento e refinamento. A Figura 66 apresenta os valores da *loss* de energia e direção. É visível que ambas apresentam uma queda gradual conforme o treinamento avança, indicando

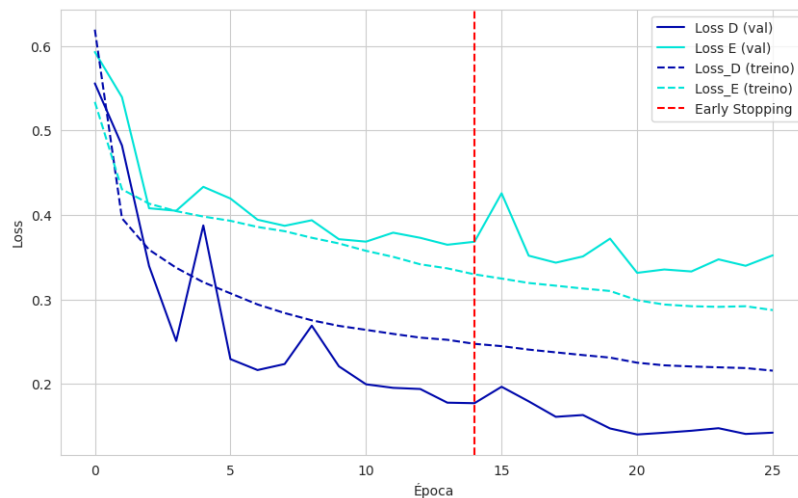
Figura 65 – Curva ROC,o modelo alcança uma AUC de 95% ainda nas fases iniciais do treinamento



Fonte: Autoria própria.

que o *backbone* convolucional está extraindo características morfológicas que são úteis para ambas as tarefas de regressão.

Figura 66 – Evolução da função de perda para as tarefas de regressão de Energia e Direção durante o primeiro estágio de treinamento.



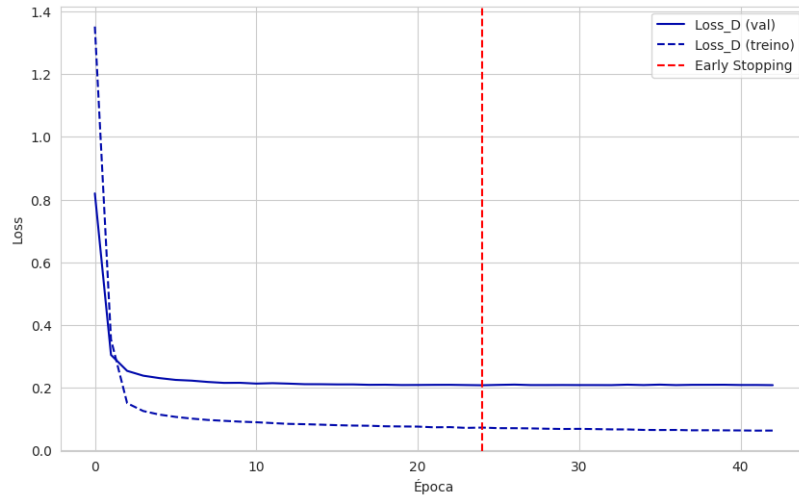
Fonte: Autoria própria.

Para a etapa seguinte os pesos do *backbone* foram travados, impedindo que a extração de features fosse prejudicada. Os pesos da *head* de Direção foram zerados e a função de perda da rede assumiu a função de perda MSE da direção, para forçar a convergência da geometria estéreo, o que aconteceu nas primeiras épocas e entrou em um platô (Figura 67).

A etapa seguinte focou no aprendizado da energia, com a adição de pesos estáticos no cálculo da função de perda e liberando a rede para *fine tuning*. A Figura 68 mostra que houve uma melhora significativa no aprendizado da regressão de energia.

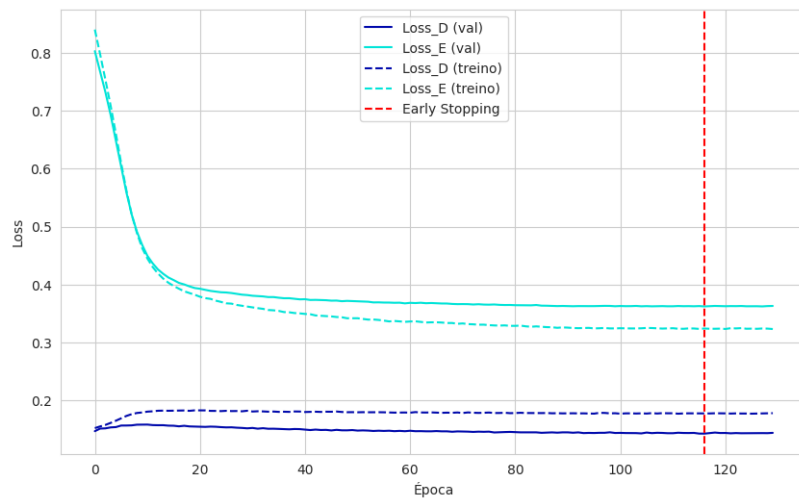
A última etapa do treinamento consistiu no *fine tuning* de todos os parâmetros da rede. Nesta fase, a introdução da camada de *Global Sum Pooling* foi essencial para a melhoria da

Figura 67 – Loss da direção, mostrando que a rede atinge a convergência da geometria estéreo



Fonte: Autoria própria.

Figura 68 – Função de perda das tarefas de regressão durante o aprendizado de energia



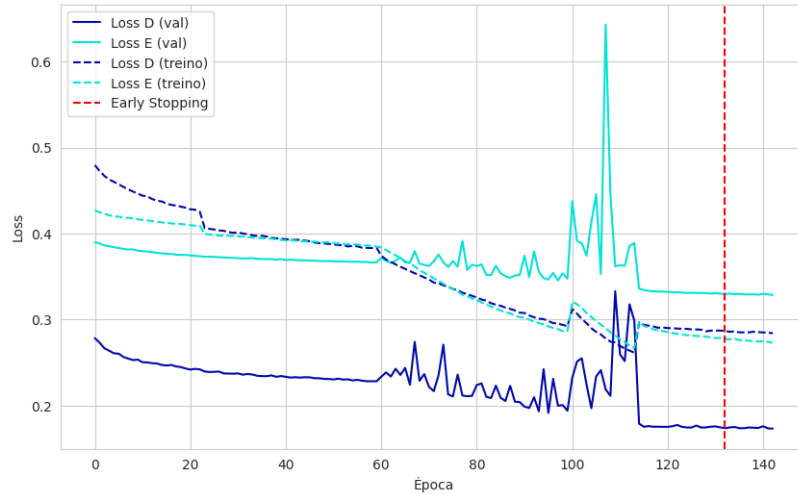
Fonte: Autoria própria.

tarefa de regressão, pois a soma preserva a informação da intensidade total de luz capturada pelo telescópio.

A Figura 69 mostra o histórico da função de perda durante a etapa final, percebe-se um comportamento atípico entre as épocas 60 e 115. Este fenômeno foi provocado por uma intervenção manual na taxa de aprendizagem que permitiu que a rede explorasse o espaço de perda em busca de uma solução mais generalista para a estereoscopia.

A queda após a restauração de uma taxa de aprendizagem menor comprova que a rede encontrou um ponto de equilíbrio entre calorimetria e geometria estéreo, com ambas alcançando o patamar de erro mínimo histórico.

Figura 69 – Função de perda das tarefas de regressão da energia e direção ao final do treinamento, com regime de exploração da taxa de aprendizagem



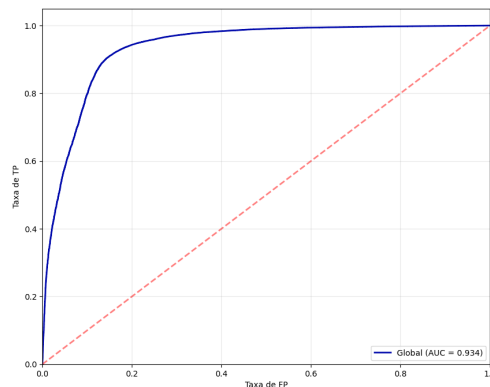
Fonte: Autoria própria.

4.2 Classificação Gama-Hádron: Global

A avaliação primária do modelo foi realizada em um cenário global, integrando eventos simulados em toda a faixa de energia de sensibilidade dos LSTs e regiões limiares de mais baixas e altas energias, e abrangendo todas as geometrias de observação (20° , 40° e 60°). Este cenário representa o desafio mais generalista para a rede neural, exigindo que o classificador generalize características morfológicas invariantes à escala de energia e à profundidade atmosférica, ao mesmo tempo em que lida também com o desafio da ausência de multiplicidade em vários eventos.

A Figura 70 apresenta a Curva ROC deste cenário. O comportamento assintótico da curva em direção ao canto superior esquerdo confirma o alto poder de discriminação do modelo. O valor de AUC de 0,934 supera referências clássicas baseadas em parâmetros de Hillas (NIETO *et al.*, 2017), validando a capacidade do *Backbone* convolucional em extrair correlações sutis que escapam à parametrização elíptica simples.

Figura 70 – Curva ROC para o cenário global, alcançando AUC de 0,934

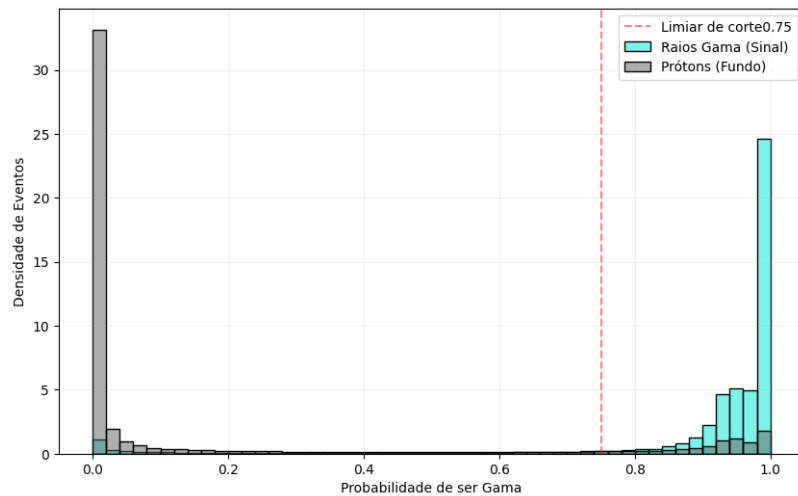


Fonte: Autoria própria.

4.3 Matriz confusão e limiar de decisão: Global

Para a extração de resultados científicos, a escolha do limiar de decisão (limiar de *gammaness*) é de extrema importância. Neste trabalho, definiu-se o limiar de 0,75 como ponto de corte operacional. Esta escolha fundamenta-se na otimização da sensibilidade da rede, buscando o compromisso ideal entre a eficiência de detecção de fótons gama e a taxa de rejeição do fundo hadrônico. Permitindo que a rede opere em uma região de alta pureza mantendo uma boa significância estatística.

Figura 71 – Limiar de decisão (Gammaness) em azul, prótons em cinza



Fonte: Autoria própria.

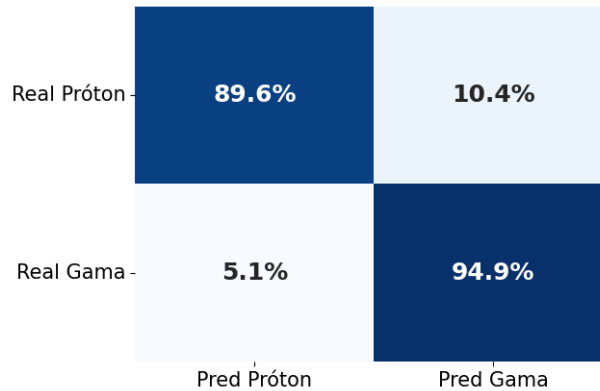
Após uma avaliação dos eventos iniciados por prótons que foram erroneamente identificados como raios gama, notou-se uma distribuição bimodal, com picos em baixas energias, onde temos poucos fótons para classificação e em uma região em torno de 100 GeV, que pode ser explicada pela dominância do decaimento do π^0 , que decai instantaneamente em $\gamma\gamma$ que mimetizam o gama primário, sendo, portanto, efeitos físicos intrínsecos à natureza do evento.

Na Figura 72 é apresentada a matriz de confusão global, demonstrando que, ao adotar o limiar de 0,75, a rede apresenta grande capacidade de rejeição do fundo hadrônico, filtrando cerca de 90% do ruído e com uma eficiência de detecção de gama de $\sim 95\%$.

O modelo demonstrou uma separação nítida entre as populações de chuviros eletromagnéticos e hadrônicos, a Acurácia, que representa a fração total de predições corretas do modelo, obtida a partir dos dados da matriz de confusão, atinge aproximadamente 92,2%. Este valor é particularmente significativo considerando toda a faixa energética de 5 a 300 GeV. Ele indica que, mesmo próximo ao limiar de detecção onde a imagem é bastante tênue e composta por poucos fotoelétrons, a rede se mantém consistente na identificação do sinal.

Enquanto a acurácia fornece uma visão geral, o *F1-Score* é a métrica utilizada para avaliar o balanço entre a Precisão (pureza da amostra) e o *Recall* (eficiência de detecção), a rede proposta alcançou um *F1-Score* de 92,4%, demonstrando a capacidade de fornecer uma amostra de dados com alta pureza para a análise científica.

Figura 72 – Matriz de confusão para o cenário global



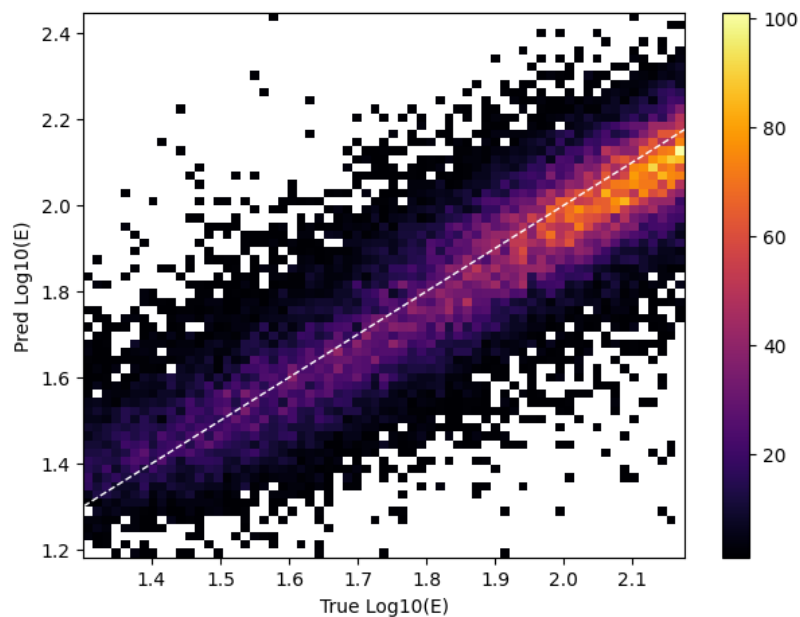
Fonte: Autoria própria.

4.4 Reconstrução de Energia: Global

A regressão de energia na rede baseia-se no princípio da calorimetria. Assumindo que a energia do fóton primário é proporcional à quantidade total de luz Cherenkov coletada, evidenciando a importância do uso do *sum pooling*. Porém essa relação tende a ser prejudicada pela distância de impacto do chuveiro e pelo ângulo de observação. Ao trabalhar com eventos nos limiares de detecção do instrumento a reconstrução torna-se suscetível a flutuações estatísticas e efeitos de seleção sistemáticos.

A Figura 73 apresenta a matriz de migração de energia para o cenário global, evidenciando a correlação entre a energia real (E_{true}) e a energia reconstruída pela rede (E_{pred}).

Figura 73 – Matriz de migração de energia para o cenário global



Fonte: Autoria própria.

Em um cenário global a rede alcança uma resolução de energia E_{res} de 30,79% com um viés de -11,41%. Este valor relativamente elevado para a resolução deve-se à degeneres-

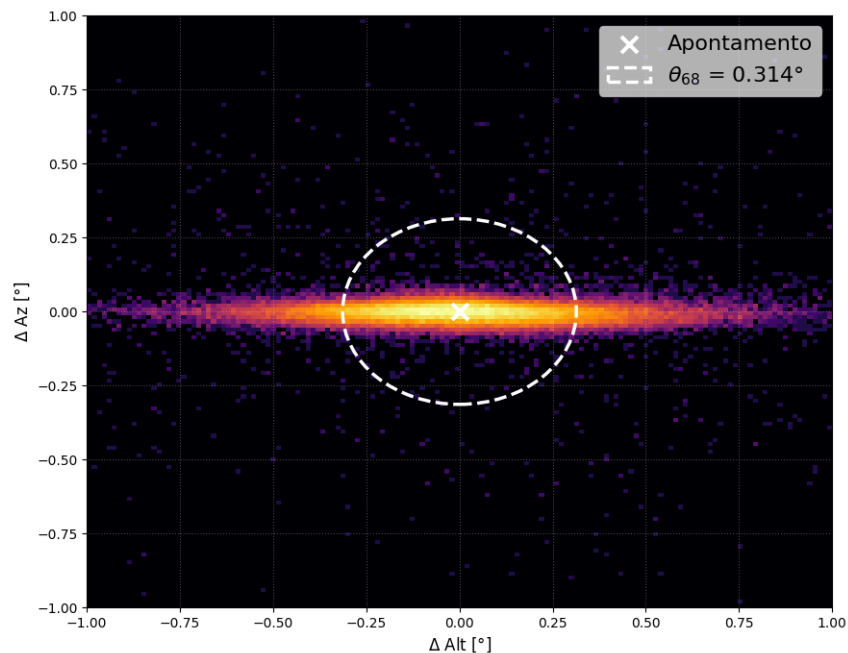
cência inerente a uma visão monocular do evento: sem a informação da distância de impacto, a rede tem dificuldade em distinguir um chuvaire de baixa energia próximo ao telescópio de um chuvaire de alta energia distante.

4.5 Reconstrução da Direção: Global

A reconstrução da direção de chegada é um dos desafios mais complexos para a arquitetura, exigindo que o modelo internalize a geometria projetiva do arranjo de telescópios em uma combinação estereoscópica, tendo dificuldade para reconstruir eventos observados por apenas um telescópio. Os resultados de resolução angular apresentados foram obtidos após a remoção do bias sistemático residual. Permitindo uma avaliação da capacidade da rede na tarefa de regressão independentemente do deslocamento da apontamento do telescópio.

A Figura 74 mostra a resolução angular θ_{68} alcançada pelo modelo, com um raio de contenção de 68% dos eventos de $0,314^\circ$. A capacidade da rede em manter um θ_{68} abaixo de $0,5^\circ$ sem a aplicação de cortes de seleção rigorosos demonstra que o modelo aprendeu a extrair a informação direcional mesmo em condições de sinal não ideal.

Figura 74 – Resolução Angular (Região de 68% de contenção) do cenário global



Fonte: Autoria própria.

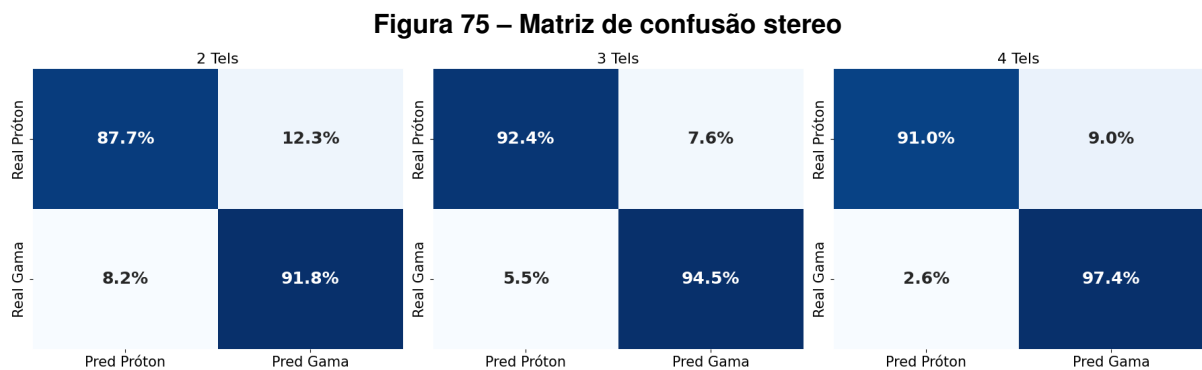
4.6 Análise de Dependências: Multiplicidade, Energia e Ângulo Zenital

Para compreender a fundo os limites operacionais da rede proposta, esta seção decompõe o desempenho global em variáveis controladas. Esta análise é fundamental para validar se

a rede aprendeu a física do detector ou se está apenas correlacionando brilho com energia de forma superficial.

4.6.1 Multiplicidade

A multiplicidade de um evento é definida pelo número de telescópios que detectaram o chuva simultaneamente. A transição do cenário mono para a estereoscopia revela o ganho real da arquitetura proporcionados pelo mecanismo de atenção. As matrizes de confusão apresentadas na Figura 75 indicam que uma visão estereoscópica melhora a tarefa de classificação.



Fonte: Autoria própria.

A Tabela 1 apresenta a comparação da classificação global segmentada pela multiplicidade. É possível notar uma melhora expressiva em todas as métricas: A rede alcança uma AUC de 94,20% e um $F1-Score$ de 94,38 %, uma resolução de energia de 19,05% e um θ_{68} de 0,197° no cenário onde o evento foi detectado por 4 telescópios. Demonstrando que ao cruzar informações de múltiplos pontos de vista, o mecanismo de atenção da rede realiza uma triangulação implícita do núcleo do chuva.

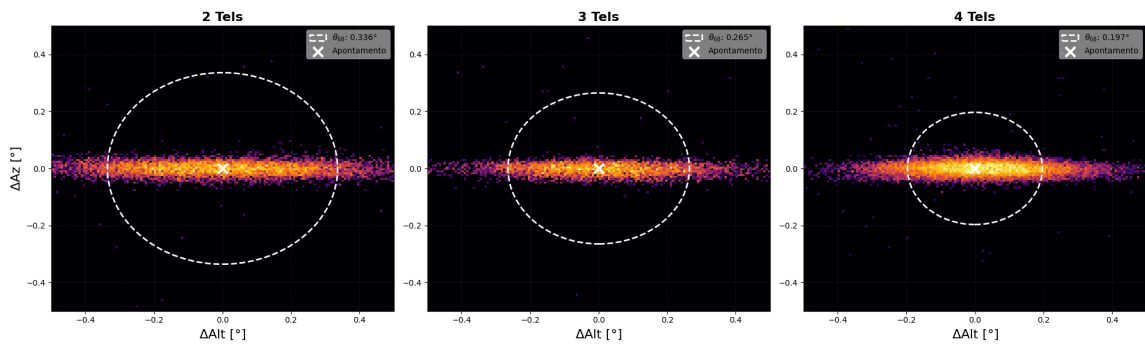
Tabela 1 – Comparação de Desempenho entre Cenários de Multiplicidade

Cenário	AUC	Acc (%)	F1-Score (%)	E_{Res} (%)	θ_{68} (°)
2 Telescópios	0.942	89.75	89.95	33.52	0.336
3 Telescópios	0.968	93.51	93,51	27.02	0.265
4 Telescópios	0.976	94.20%	94.38	22.31	0.197

A Figura 76 apresenta a resolução angular para observações estereoscópicas onde 2, 3 ou 4 telescópios foram acionados. É possível perceber a diminuição no valor de θ_{68} , comprovando que a triangulação é um fator fundamental para a reconstrução da direção de chegada da partícula incidente.

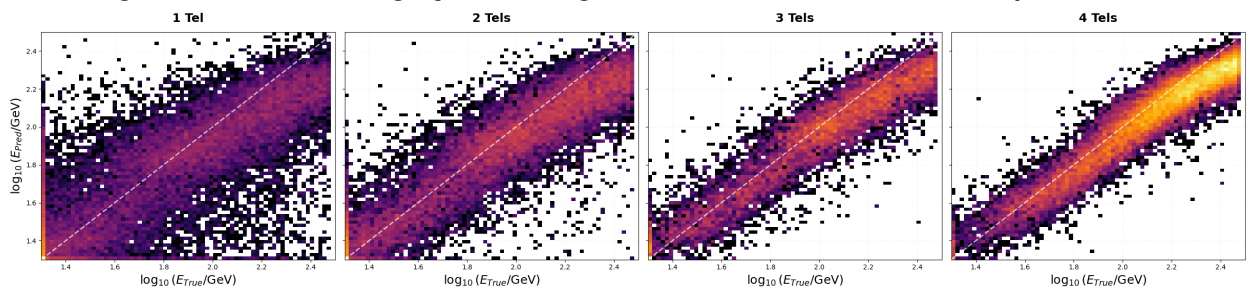
Um resultado parecido aparece ao olharmos para a matriz de migração de energia apresentada na Figura 77, havendo uma maior concentração de eventos ao redor da linha diagonal no cenário com 4 telescópios.

Figura 76 – Resolução Angular em um cenário de estereoscopia para 2, 3 e 4 telescópios



Fonte: Autoria própria.

Figura 77 – Matriz de migração de energia em diferentes cenários de multiplicidade



Fonte: Autoria própria.

4.6.2 Energia

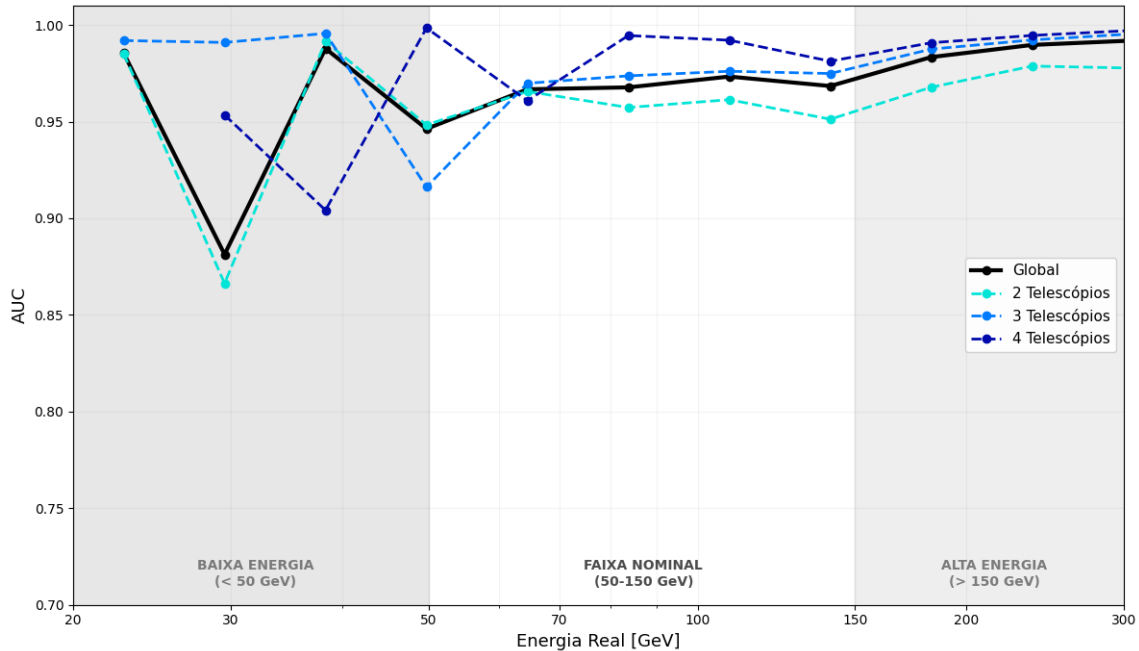
A análise da performance da rede foi estratificada em três domínios energéticos distintos para avaliar a resposta do modelo frente aos desafios instrumentais específicos de cada faixa.

- Baixa energia - até 50 GeV: Focada na análise do comportamento próximo ao limiar de gatilho, onde a relação sinal-ruído é crítica.
- Nominal - 50 GeV a 150 GeV: Correspondente à janela de sensibilidade prioritária do LST, onde o telescópio atinge o máximo de eficiência para o qual foi projetado.
- Alta energia - acima de 150 GeV: Esta faixa de energia se sobrepõe com a faixa de observação dos MSTs, os chuveiros são mais brilhantes e tem uma morfologia mais bem definida.

A Figura 78 apresenta aos valores da AUC obtida em diferentes níveis de energia e multiplicidade. Percebe-se que a rede apresenta maior dificuldade de classificação em baixa energia, onde o número de fótons que atingem o espelho do telescópio é muito baixo, limitando a definição morfológica das imagens dos chuveiros. Ao atingir a faixa de energia da região nominal de operação, a performance estabiliza com valores de AUC superiores a 0,95. O melhor desempenho é alcançado em situação de estereoscopia total (evento detectados pelos 4 telescópios). A partir de 150 GeV há uma melhor definição morfológica devido à maior quantidade

de fótons incidentes, nesse cenário a rede atinge AUC superiores a 0,98 em praticamente todas as configurações de observação estereoscópicas.

Figura 78 – Desempenho da classificação (AUC) por energia em cenários com estereoscopia



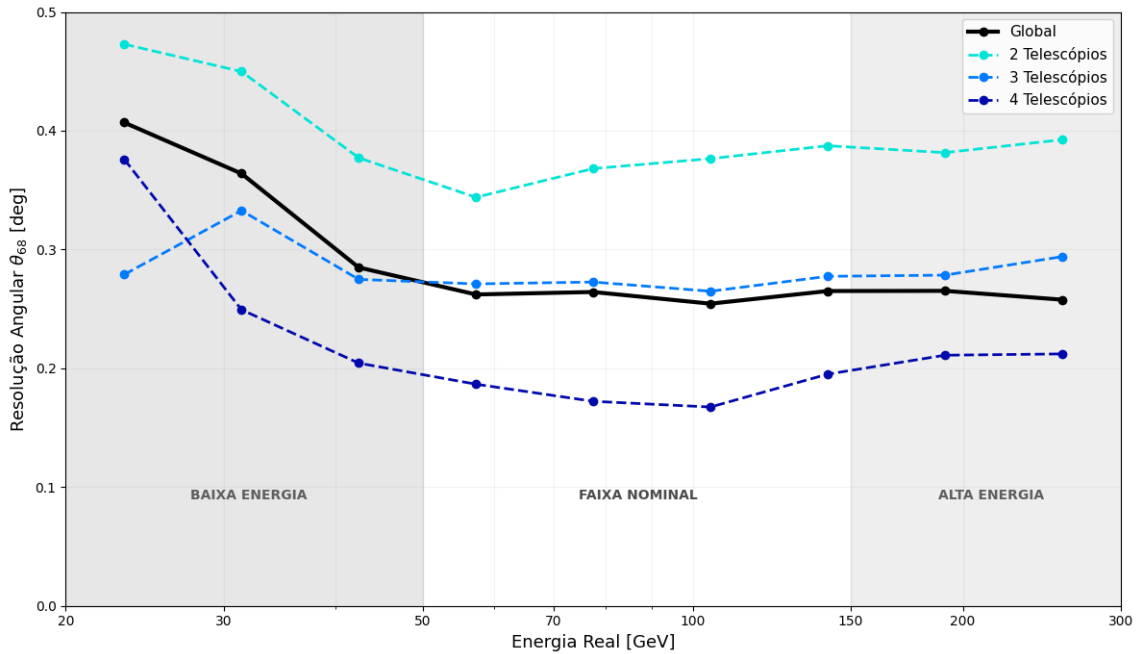
Fonte: Autoria própria.

A Figura 79 ilustra a evolução da resolução angular (θ_{68}) em relação à energia real dos eventos para diferentes níveis de multiplicidade. Observa-se uma performance inferior na região de Baixa Energia devido à dificuldade de reconstruir com precisão chuveiros que geram menos luz Cherenkov. Novamente fica evidenciado o desempenho mais alto quando há observação estereoscópica total, com a observação com 4 telescópios mantendo valores de θ_{68} abaixo de $0,2^\circ$ em toda a faixa nominal. Em contrapartida, a configuração com apenas 2 telescópios mostra uma performance inferior, reforçando que o aumento da informação geométrica é fundamental para mitigar erros de reconstrução da direção da partícula incidente.

4.6.3 Ângulo Zenital

A influência do ângulo zenital na performance do modelo para o modo estéreo (≥ 2 telescópios) é detalhada na Tabela 2. Observa-se que, mesmo com a variação angular, a AUC se mantém acima de 0,93. Obtendo melhor performance em 60° , onde também atinge melhor resolução de energia (25,77%). Fisicamente, esse comportamento pode ser atribuído ao fato de que, em ângulos mais inclinados, o chuveiro atmosférico se desenvolve a uma distância maior dos telescópios, permitindo que o sistema de detecção capture uma visão mais completa da cascata de partículas, o que favorece a discriminação morfológica e a reconstrução energética. Já para a reconstrução da direção de chegada da partícula, a região de maior precisão se dá na faixa de energia entre 50 a 150 GeV em observações a 40° , atingindo um θ_{68} de $0,205^\circ$.

Figura 79 – Resolução Angular θ_{68} por níveis de energia em cenários com estereoscopia



Fonte: Autoria própria.

Este valor representa uma melhoria significativa em relação à média global, evidenciando que a geometria estéreo é otimizada para fontes observadas em inclinações moderadas. Na região de Alta Energia (≥ 150 GeV), a performance de classificação atinge patamares de excelência, com AUC de 0,990 em 40° , reforçando a robustez da rede em identificar fótons gama de alta energia independentemente da inclinação da observação.

Tabela 2 – Desempenho por ângulo zenital e faixa de energia em observação estereoscópica (≥ 2 Tels).

Zênite	Cenário	AUC	E_{Res} (%)	θ_{68} ($^\circ$)
20°	Total (20°)	0,952	26,62	0,326
	Baixa E (< 50 GeV)	0,933	31,23	0,416
	Faixa Ótima (50-150 GeV)	0,955	23,44	0,265
	Alta E (≥ 150 GeV)	0,980	27,75	0,301
40°	Total (40°)	0,970	26,00	0,253
	Baixa E (< 50 GeV)	0,969	27,95	0,346
	Faixa Ótima (50-150 GeV)	0,977	23,25	0,205
	Alta E (≥ 150 GeV)	0,990	27,92	0,229
60°	Total (60°)	0,978	25,77	0,290
	Baixa E (< 50 GeV)	0,982	106,18	0,383
	Faixa Ótima (50-150 GeV)	0,967	22,16	0,331
	Alta E (≥ 150 GeV)	0,988	27,47	0,267

4.6.4 Gama pontual vs Gama Difuso

A análise comparativa entre as componentes de emissão é apresentada na Tabela 3.

Tabela 3 – Comparativo de desempenho entre as componentes pontuais e difusas (4 Tels, 40°, 50-150 GeV).

Componente	AUC	E_{Res} (%)	θ_{68} (°)
Gama Pontual	0,991	19,25	0,126
Gama Difuso	0,918	21,37	0,905

Observa-se que a capacidade de discriminação (AUC) para o componente pontual é superior à do componente difuso. No cenário de estereoscopia total, a componente pontual atinge uma AUC de 0,991 na faixa de energia nominal (50 – 150 GeV) para observações em 40°, enquanto a componente difusa alcança 0,918 no mesmo regime.

Em contrapartida, a reconstrução energética se mantém estável, apresentando resoluções de 21,37% para o componente difuso e 19,25% para o pontual na mesma faixa de energia a 40°. A maior divergência ocorre na reconstrução direcional: para fótons provenientes de fontes pontuais, a resolução angular (θ_{68}) atinge a marca de 0,128°, enquanto que em emissões difusas esse valor aumenta para 0,958°.

5 CONCLUSÃO

O *Cherenkov Telescope Array Observatory* representa o expoente da nova geração de observatórios dedicados ao estudo de raios gama, prometendo uma sensibilidade uma ordem de grandeza superior aos instrumentos atuais. No entanto, o volume de dados sem precedentes e a necessidade de maximizar a sensibilidade na faixa de baixa energia impõem desafios que os métodos tradicionais de reconstrução enfrentam com dificuldade.

Nesta dissertação, propusemos e validamos uma arquitetura de *Deep Learning* multi-tarefa baseada em redes siamesas, projetada especificamente para o arranjo de Telescópios de Grande Porte (LSTs). A abordagem rompe com a parametrização clássica de Hillas, aprendendo a extrair informações diretamente das imagens brutas.

5.1 Principais Resultados

A análise dos resultados, fundamentada em simulações de Monte Carlo, permite estabelecer três conclusões principais sobre a física da reconstrução:

Demonstramos que o treinamento simultâneo de tarefas com naturezas distintas (classificação e regressão) pode sofrer instabilidades devido à competição de gradientes. A implementação de uma estratégia de *Curriculum Learning* e Treinamento Multi-Estágio provou ser de grande importância para o resultado final. Ao forçar a rede a aprender primeiro a morfologia dos eventos (classificação) e introduzir progressivamente a complexidade geométrica (direção) e calorimétrica (energia), conseguimos evitar mínimos locais onde abordagens convencionais estagnariam.

A introdução da camada de *Global Sum Pooling* combinada com o mecanismo de atenção permitiu que a rede recuperasse a linearidade da resposta energética. Observou-se uma melhoria drástica na resolução de energia. Isso confirma que a rede aprendeu a realizar a triangulação implícita típica das observações estereoscópicas.

Ao olharmos para o desempenho em diferentes níveis de energia, identificamos que o limiar de energia inferior é onde o desempenho da rede é mais prejudicado devido ao número reduzido de fótons produzidos pelos chuveiros. A rede alcança desempenho satisfatório na região nominal de operação do LST, em eventos onde os 4 telescópios são acionados atinge-se AUC superior a 0,97 na classificação de eventos, Resolução angular abaixo de $0,2^\circ$ e Resolução de energia $< 20\%$. Já no limiar superior a rede atingiu resultados ainda mais expressivos.

5.2 Contribuições para a área

A alta eficiência de classificação e pureza das amostras e a resolução angular obtida na faixa de energia nominal dos LSTs têm implicações diretas para os *Key Science Projects* do CTAO, principalmente no que diz respeito à busca por matéria escura e GRBs de alto *redshift*.

Ao inserir o conceito de Redes Siamesas na arquitetura, este trabalho introduz uma abordagem inovadora para a reconstrução de eventos estéreos. Diferente dos métodos clássicos que dependem da parametrização geométrica de imagens individuais, a arquitetura siamesa permite que o modelo aprenda representações compartilhadas entre os telescópios. Possibilitando uma extração de características que preserva a correlação espacial entre as diferentes visões do mesmo chuveiro atmosférico, resultando em uma estimativa de direção e energia mais robusta.

5.3 Trabalhos Futuros

Apesar dos avanços, este trabalho identifica limitações que abrem caminho para pesquisas futuras. Como todo o treinamento e avaliação da rede foi realizado utilizando dados de simulações Monte Carlo, o desempenho do modelo quando aplicado em dados reais pode sofrer variações devido à discrepâncias na resposta do instrumento, ou variáveis físicas, como as condições da atmosfera local, portanto o uso de técnicas de adaptação de domínio podem ser de grande valia para que o desempenho em dados reais

Este trabalho limitou a energia até pouco acima do limiar de energia dos LSTs, onde a operação conjunta com MSTs pode ser um trunfo estereoscópico. Embora a arquitetura proposta funcione bem para um número fixo de telescópios idênticos, para outras configurações pode ser interessante a migração para redes neurais de grafos, que podem lidar nativamente com a geometria irregular e variável do arranjo, sem que o custo computacional escale exponencialmente.

O alto desempenho da rede proposta ao tratar raios gama difusos sugere que existe a possibilidade de maximizar o campo de visão através de apontamento divergente, permitindo *surveys* de todo o céu de forma mais rápida, bem como para a busca da posição de transientes e observação de fontes extensas.

REFERÊNCIAS

- AAB, A. *et al.* Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} eV. **Science**, American Association for the Advancement of Science (AAAS), v. 357, n. 6357, p. 1266–1270, set. 2017. ISSN 1095-9203. Disponível em: <http://dx.doi.org/10.1126/science.aan4338>.
- ACHARYA, B. *et al.* **Science with the Cherenkov Telescope Array**. WORLD SCIENTIFIC, 2018. ISBN 9789813270091. Disponível em: <http://dx.doi.org/10.1142/10986>.
- AHARONIAN, F. A. **Very High Energy Cosmic Gamma Radiation**. WORLD SCIENTIFIC, 2004. Disponível em: <https://www.worldscientific.com/doi/abs/10.1142/4657>.
- AMBROSI, G. *et al.* **The Cherenkov Telescope Array Large Size Telescope**. 2013. Disponível em: <https://arxiv.org/abs/1307.4565>.
- ARBELETICHE, L. B. **Simulation of extensive air showers and its detection**. 2021. 166 p. Tese (Tese (Doutorado)) — Instituto de Física de São Carlos - Universidade de São Paulo, São Paulo, 2021.
- BATISTA, R. A. **The Quest for the Origins of Ultra-High-Energy Cosmic Rays**. 2025. Disponível em: <https://arxiv.org/abs/2412.17201>.
- BENGIO, Y. *et al.* Curriculum learning. *In: Proceedings of the 26th Annual International Conference on Machine Learning*. New York, NY, USA: Association for Computing Machinery, 2009. (ICML '09), p. 41–48. ISBN 9781605585161. Disponível em: <https://doi.org/10.1145/1553374.1553380>.
- BERNLÖHR, K. **Measuring cosmic-ray and gamma-ray air showers**. 2007. <https://www.mpi-hd.mpg.de/hfm/CosmicRay/ShowerDetection.html>. Acessado em: 15 out. 2025.
- BERNLÖHR, K. Simulation of imaging atmospheric cherenkov telescopes with corsika and sim_telarray. **Astroparticle Physics**, Elsevier BV, v. 30, n. 3, p. 149–158, out. 2008. ISSN 0927-6505. Disponível em: <http://dx.doi.org/10.1016/j.astropartphys.2008.07.009>.
- Blackett, P. M. S. A possible contribution to the night sky from the Cerenkov radiation emitted by cosmic rays. *In: The Emission Spectra of the Night Sky and Aurorae*. [S.l.: s.n.], 1948. p. 34.
- BREISKY, B. **When Victor Hess Discovered Cosmic Rays in a Hydrogen Balloon**. 2012. Acessado em: 15 out. 2025. Disponível em: <https://www.nytimes.com/2012/08/07/science/space/when-victor-hess-discovered-cosmic-rays-in-a-hydrogen-balloon.html>.
- CARUANA, R. Multitask learning. **Machine Learning**, v. 28, 07 1997.
- CHALMERS, M. Ascent commemorates cosmic-ray pioneers. **CERN Courier**, mar 2020. Disponível em: <https://cerncourier.com/a/ascent-commemorates-cosmic-ray-pioneers/>.
- CHERENKOV, P. A. Visible luminescence of pure liquids under the influence of γ -radiation. **Dokl. Akad. Nauk SSSR**, v. 2, n. 8, p. 451–454, 1934.
- Cherenkov Telescope Array Observatory. **Data and Computing**. <https://www.ctao.org/emission-to-discovery/data-and-computing/>. Acessado em: 01 nov. 2025. Disponível em: <https://www.ctao.org/emission-to-discovery/data-and-computing/>.

CORREIA, C. E. C.; NOGUEIRA, R. de S. **Redes Neurais Artificiais Aplicadas no Monitoramento de Ativos em Subestações de Energia Elétrica**. 2023. Trabalho de Conclusão de Curso (Graduação).

CORTINA, J. **Status of the Large Size Telescopes of the Cherenkov Telescope Array**. 2019. Disponível em: <https://arxiv.org/abs/1907.10146>.

CTA Observatory. **CTAO Performance**. 2021. Disponível em: <https://www.cta-observatory.org/wp-content/uploads/2021/08/CTA-Performance-prod5-v0.1-North-NorthSouth50h-DifferentialSensitivity-OtherInstruments.png>. Acesso em: 07 jan. 2024.

CTA Observatory. **Array Sites**. 2024. Accessed: 2024-05-25. Disponível em: <https://www.ctao.org/emission-to-discovery/array-sites/>.

CTA Observatory. **DL1 Data Handler**. 2024. https://github.com/cta-observatory/dl1-data-handler/blob/master/notebooks/test_image_mapper.ipynb. Accessed: 2024-08-25.

CTA Observatory. **pyeventio: Python library to read eventio data**. 2024. <https://github.com/cta-observatory/pyeventio>. Acesso em: 13 de set. de 2024.

DORMAN, I.; DORMAN, L. How cosmic rays were discovered and why they received this misnomer. **Advances in Space Research**, v. 53, n. 10, p. 1388–1404, 2014. ISSN 0273-1177. Cosmic Ray Origins: Viktor Hess Centennial Anniversary. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0273117713002500>.

FILIPOVIĆ, M. D.; PAYNE, J. L.; TOTHILL, N. F. H. Multimessenger astronomy in practice: Celestial sources in action. *In: Multimessenger Astronomy in Practice*. IOP Publishing, 2021, (2514-3433). p. 1–1 to 1–27. ISBN 978-0-7503-2344-4. Disponível em: <https://dx.doi.org/10.1088/2514-3433/ac2256ch1>.

FRANK, I. M.; TAMM, I. Y. Coherent visible radiation of fast electrons passing through matter. **Physics-Uspekhi**, v. 93, p. 388–393, 1937. Disponível em: <https://api.semanticscholar.org/CorpusID:120903501>.

GAMOW, G. **The Great Physicists from Galileo to Einstein**. New York: Dover Publications, 1981. Originalmente publicado em 1961 como "Biography of Physics". ISBN 9780486136813.

GILER, A. G. D. **Measuring the shower maximum of extensive air showers using imaging atmospheric Cherenkov telescopes**. 2019. Dissertação (Dissertação (Mestrado em Física)) — Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019. Disponível em: <https://doi.org/10.11606/D.76.2019.tde-20092019-110630>.

GOCKEL, A. Über die in der atmosphäre vorhandene durchdringende strahlung. **Physikalische Zeitschrift**, v. 10, p. 845–847, 1909.

GOCKEL, A. Lufterlektrische beobachtungen bei einer ballonfahrt. **Physikalische Zeitschrift**, v. 11, p. 280–282, 1910.

GOCKEL, A. Messungen der durchdringenden strahlung bei ballonfahrten. **Physikalische Zeitschrift**, v. 12, p. 595–597, 1911.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

GRUPEN, C. **Astroparticle Physics**. [S.l.]: Springer Cham, 2020. ISBN 9783030273415.

- GU, J. *et al.* **Recent Advances in Convolutional Neural Networks**. 2017. Disponível em: <https://arxiv.org/abs/1512.07108>.
- HECK, D. *et al.* **CORSIKA: a Monte Carlo code to simulate extensive air showers**. 1998.
- HILLAS, A. M. Cerenkov Light Images of EAS Produced by Primary Gamma Rays and by Nuclei. *In*: Jones, F. C. (Ed.). **19th International Cosmic Ray Conference (ICRC19), Volume 3**. [S.l.: s.n.], 1985. (International Cosmic Ray Conference, v. 3), p. 445.
- HOFMANN, W.; ZANIN, R. The cherenkov telescope array. *In*: _____. **Handbook of X-ray and Gamma-ray Astrophysics**. Singapore: Springer Nature Singapore, 2024. p. 2787–2833. ISBN 978-981-19-6960-7.
- HORVATH, J. E. **High-Energy Astrophysics. A Primer**. [S.l.]: Springer, 2022. (Undergraduate Lecture Notes in Physics). ISBN 978-3-030-92158-3, 978-3-030-92161-3, 978-3-030-92159-0.
- JACQUEMONT, M. *et al.* First Full-Event Reconstruction from Imaging Atmospheric Cherenkov Telescope Real Data with Deep Learning. *In*: **2021 International Conference on Content-Based Multimedia Indexing (CBMI)**. Lille, France: IEEE, 2021. p. 6 p. Disponível em: <https://hal.science/hal-03241475>.
- KENDALL, A.; GAL, Y.; CIPOLLA, R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. **CoRR**, abs/1705.07115, 2017. Disponível em: <http://arxiv.org/abs/1705.07115>.
- KRYUKOV, A. P.; DEMICHEV, A. P.; ILYIN, V. A. Machine learning in gamma astronomy. **Moscow University Physics Bulletin**, Allerton Press, v. 79, n. S2, p. S690–S699, dez. 2024. ISSN 1934-8460. Disponível em: <http://dx.doi.org/10.3103/S002713492470214X>.
- LATTES, C. M. G. *et al.* Processes involving charged mesons. **Nature**, v. 159, n. 4047, p. 694–697, 1947.
- LEE, C. M. G. **O espaço ROC para um classificador "melhor" e um classificador "pior"**. 2018. Wikimedia Commons, via Wikipedia. Acessado em: 25 de out. de 2025. Disponível em: https://en.wikipedia.org/wiki/Receiver_operating_characteristic#/media/File:Roc_curve.svg.
- LONGAIR, M. S. **High Energy Astrophysics**. 3. ed. [S.l.]: Cambridge University Press, 2011.
- MIRZOYAN, R. The development of ground-based gamma-ray astronomy: A historical overview of the pioneering experiments. *In*: _____. **Handbook of X-ray and Gamma-ray Astrophysics**. Singapore: Springer Nature Singapore, 2022. p. 1–26. ISBN 978-981-16-4544-0. Disponível em: https://doi.org/10.1007/978-981-16-4544-0_62-1.
- NIELSEN, M. A. **Neural Networks and Deep Learning**. [S.l.]: Determination Press, 2015.
- NIETO, D. *et al.* **Studying deep convolutional neural networks with hexagonal lattices for imaging atmospheric Cherenkov telescope event reconstruction**. 2019. Disponível em: <https://arxiv.org/abs/1912.09898>.
- NIETO, D. *et al.* **Exploring deep learning as an event classification method for the Cherenkov Telescope Array**. 2017. Disponível em: <https://arxiv.org/abs/1709.05889>.
- NORRIS, R. Data Science (Mining and Processing) with Machine Learning in the Era of Multimessenger Astronomy. *In*: Filipović, M. D.; Tothill, N. F. H. (Ed.). **Multimessenger Astronomy in Practice: Celestial Sources in Action**. [S.l.: s.n.], 2021. p. 12–1.
- OBSERVATORY, C. **CTAO Simulation Telescope Models for CORSIKA and sim_telarray - Prod6**. Zenodo, 2024. Disponível em: <https://doi.org/10.5281/zenodo.14198379>.

- PARSONS, R.; HINTON, J. A monte carlo template based analysis for air-cherenkov arrays. **Astroparticle Physics**, Elsevier BV, v. 56, p. 26–34, abr. 2014. ISSN 0927-6505. Disponível em: <http://dx.doi.org/10.1016/j.astropartphys.2014.03.002>.
- PEREIRA, J. P. *et al.* **Binary coalescences as sources of Ultra-High Energy Cosmic Rays**. 2024. Disponível em: <https://arxiv.org/abs/2307.06200>.
- PRINCE, S. **Understanding Deep Learning**. MIT Press, 2023. ISBN 9780262377102. Disponível em: <https://books.google.com.br/books?id=pPyxEAAAQBAJ>.
- PTITSYNA, K. V.; TROITSKY, S. V. Physical conditions in potential accelerators of ultra-high-energy cosmic rays: updated hillas plot and radiation-loss constraints. **Physics-Uspekhi**, v. 53, n. 7, p. 691, oct 2010. Disponível em: <https://doi.org/10.3367/UFNe.0180.201007c.0723>.
- RASCHKA, S.; MIRJALILI, V. **Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow**. 2nd. ed. UK: Packt, 2017.
- SANT'ANNA JUNIOR, A. B. d. **Análise de chuviros atmosféricos extensos: física de partículas com aceleradores extraterrestres**. 2022. Dissertação (Dissertação (Mestrado em Física)) — Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022. Disponível em: <https://doi.org/10.11606/D.76.2022.tde-06042022-094316>.
- SCHMIDT, F.; KNAPP, J. **CORSIKA Shower Images**. 2005. Site pessoal. Disponível em: <https://www-zeuthen.desy.de/~jknapp/fs/showerimages.html>. Acesso em: 25 mai. 2024.
- SHARMA, M. *et al.* Gamma/hadron segregation for a ground based imaging atmospheric cherenkov telescope using machine learning methods: Random forest leads. **Research in Astronomy and Astrophysics**, IOP Publishing, v. 14, n. 11, p. 1491–1503, out. 2014. ISSN 1674-4527. Disponível em: <http://dx.doi.org/10.1088/1674-4527/14/11/012>.
- SHILON, I. *et al.* Application of deep learning methods to analysis of imaging atmospheric cherenkov telescopes data. **Astroparticle Physics**, v. 105, p. 44–53, 2019. ISSN 0927-6505. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0927650518301178>.
- SILVA, D. J. **Classificação de Chuviros Atmosféricos Utilizando Redes Neurais Convolucionais**. 2024. Trabalho de Conclusão de Curso (Graduação).
- SILVA, D. J.; MELLO, A. J. T. S. Convolutional neural networks for gamma and hadron event classification. **Boletim da Sociedade Astronômica Brasileira**, v. 36, n. 1, p. 58–60, 2025.
- SOUTO, B. F. **Configurações para um arranjo de telescópios Cherenkov por otimização evolutiva multiobjetivo**. nov. 2018. 260 p. Tese (Doutorado) — CBPF - Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro - RJ, nov. 2018.
- SPURIO, M. **Probes of Multimessenger Astrophysics: Charged cosmic rays, neutrinos, γ -rays and gravitational waves**. [S.l.]: Springer, 2018. (Astronomy and Astrophysics Library).
- SZELISKI, R. **Computer Vision: Algorithms and Applications**. 2nd. ed. [S.l.]: Springer Cham, 2022. XXII, 925 p. (Texts in Computer Science). ISBN 978-3-030-34374-3.
- Wakely, S. P.; Horan, D. TeVCat: An online catalog for Very High Energy Gamma-Ray Astronomy. **International Cosmic Ray Conference**, v. 3, p. 1341–1344, 2008.
- WEEKES, T. **Very High Energy Gamma-Ray Astronomy**. CRC Press, 2003. (Series in Astronomy and Astrophysics). ISBN 9781420033199. Disponível em: <https://books.google.cd/books?id=wSTOBgAAQBAJ>.

Weekes, T. C. *et al.* Observation of TeV Gamma Rays from the Crab Nebula Using the Atmospheric Cerenkov Imaging Technique. , v. 342, p. 379, jul. 1989.

WHITE, R.; BERNLOHR, K.; DESY. **Cherenkov Effect**. 2018. Online. Acessado em: 27 dez. 2025. Disponível em: <https://flic.kr/p/2cxkxLo>.

YANG, H.-B. *et al.* Application of the drs4 chip for ghz waveform digitizing circuits. **Chinese Physics C**, IOP Publishing, v. 39, n. 5, p. 056101, maio 2015. ISSN 1674-1137. Disponível em: <http://dx.doi.org/10.1088/1674-1137/39/5/056101>.

YUN, P. *et al.* Focal loss in 3d object detection. **CoRR**, abs/1809.06065, 2018. Disponível em: <http://arxiv.org/abs/1809.06065>.

ZHAO, Y. *et al.* **HexCNN: A Framework for Native Hexagonal Convolutional Neural Networks**. 2021. Disponível em: <https://arxiv.org/abs/2101.10897>.

ZYLA P A, a. B. R. M. *et al.* Review of Particle Physics. **Progress of Theoretical and Experimental Physics**, n. 8, p. 083C01, 2020. ISSN 2050-3911. Disponível em: <https://doi.org/10.1093/ptep/ptaa104>.

APÊNDICES

APÊNDICE A – Infraestrutura computacional

As simulações de Monte Carlo (CORSIKA e sim_telarray) foram executadas de forma distribuída utilizando infraestrutura computacional baseada em processadores x86_64 sob ambiente Linux.

Por sua vez, o treinamento, validação e inferência da rede neural exigiram o uso de aceleração gráfica. As especificações de hardware e ambiente de software estão detalhadas na Tabela 4.

Tabela 4 – Especificações da infraestrutura computacional e ambiente de software utilizados no treinamento da rede neural.

Componente / Software	Especificação
GPU	NVIDIA GeForce RTX 3090 (24 GB VRAM)
CPU	AMD Ryzen 7 7700
Memória RAM	62 GB
Sistema Operacional	Ubuntu 22.04.3 LTS
Framework de Deep Learning	TensorFlow 2.14.0 (API keras 2.14)
Linguagem	Python 3.11
Aceleração e Drivers	CUDA Toolkit 11.8 / cuDNN 8