

Identificación de fotones provenientes del decaimiento radiativo del bosón Z mediante Redes Neuronales Artificiales Convolucionales

• Trabajo de Diploma FCE-UNLP •
Abril 2020

Santiago Javier Noacco Rosende

Director
Dr. Fernando Monticelli

Índice

1. Motivación y objetivo de este trabajo
2. El LHC y ATLAS
 - 2.1. Subdetectores
3. Sistema de Trigger
4. Fotones en ATLAS
 - 4.1. Reconstrucción
 - 4.2. Identificación
 - 4.3. Aislamiento
5. Redes Neuronales
 - 5.1. Características del entrenamiento
 - 5.2. Redes Neuronales Convolucionales

Índice

- 6. Identificación de fotones con CNNs
 - 6.1. Almacenamiento de eventos
 - 6.2. Definición de **señal** y **fondo**
 - 6.3. Creando Imágenes de los *topo-clusters*
 - 6.4. Entrenando las CNN
- 7. Resultados
- 8. Conclusiones

Motivación

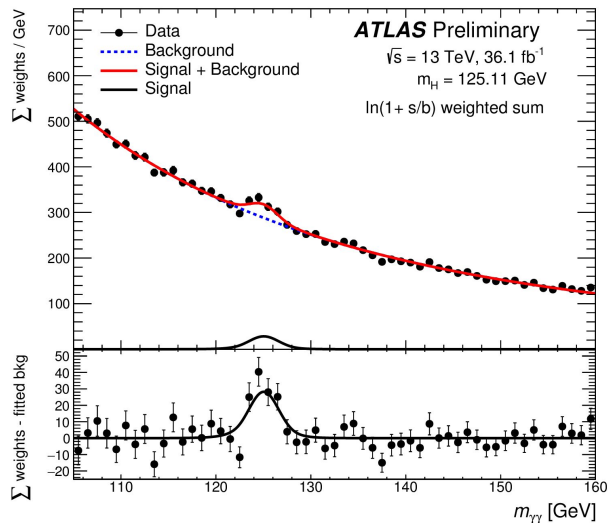
En el programa de física de ATLAS, que contempla:

- ***mediciones*** de precisión de procesos del Modelo Estándar (SM), de las propiedades del bosón de Higgs
- ***búsquedas*** de materia oscura, de evidencia de Supersimetría (SUSY), de física más allá del SM

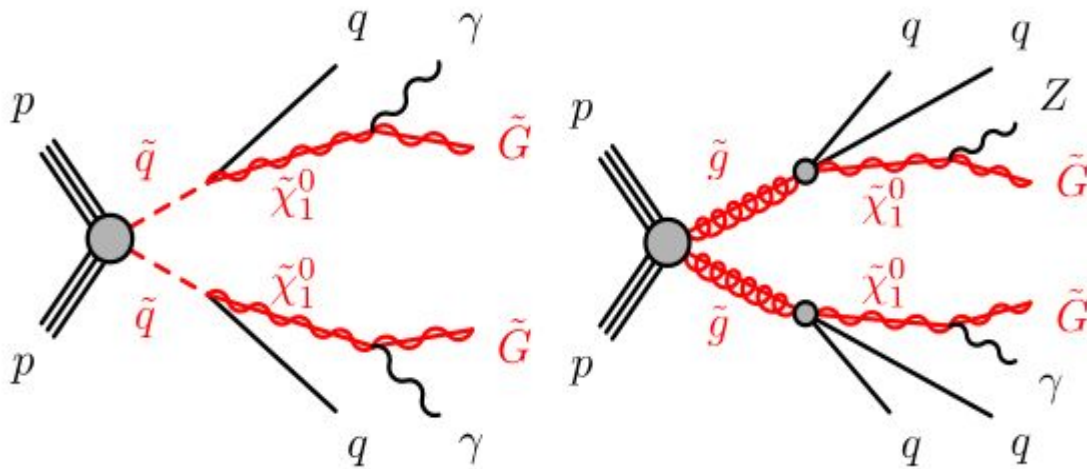
La física involucrada en varias de estas búsquedas y mediciones implican procesos con fotones en el estado final.

Es por ello que resulta fundamental contar con un ***método eficiente en la identificación de fotones y separación de fondo a partir de los datos de ATLAS.***

Motivación



Espectro de masa invariante de dos fotones seleccionados en mediciones del decaimiento del Higgs. La línea roja es el ajuste a los datos donde el modelo ajustado incluye la forma esperada para el bosón de Higgs como también el fondo.



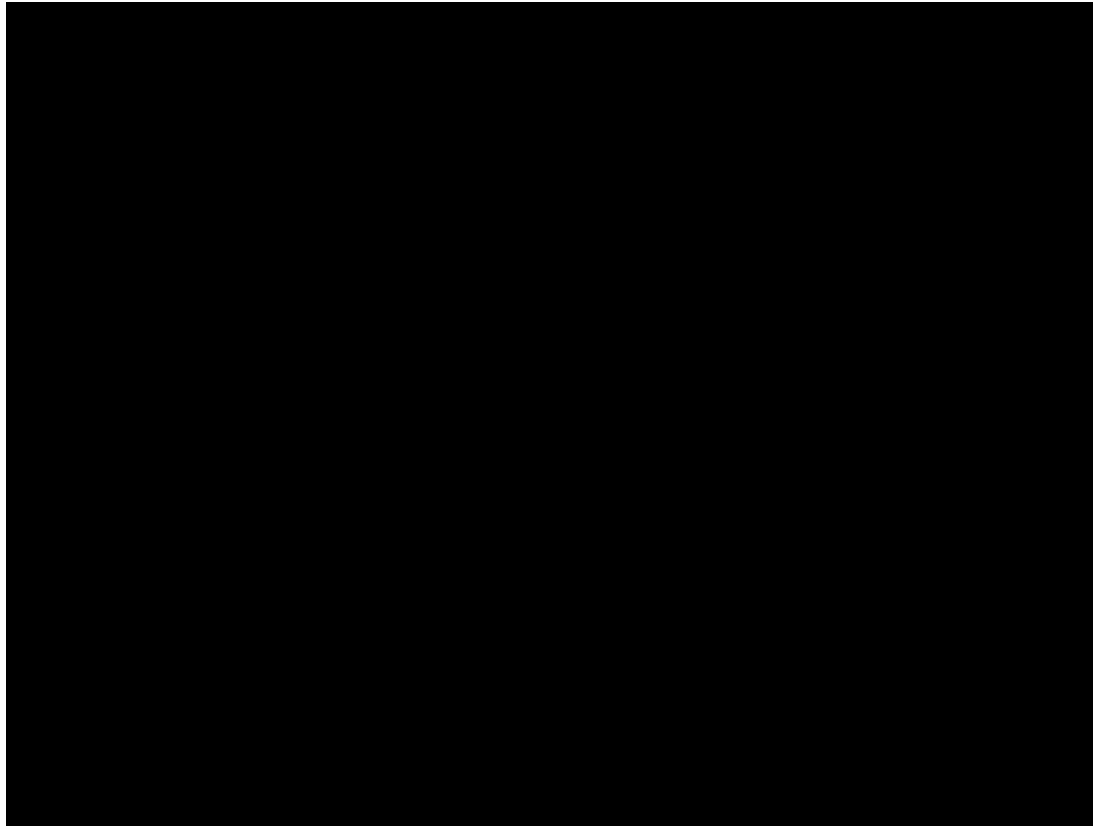
Procesos típicos de producción y decaimiento de squarks (centro) y gluinos (derecha)

Objetivo

- ***Desarrollar un método*** alternativo al actual (basado en cortes rectangulares en el EMC) ***de identificación de fotones por medio del uso de CNN.***
- ***Evaluar la eficiencia en la identificación de fotones derivados del decaimiento radiativo del bosón Z*** entre el método vigente que emplea cortes rectangulares en el EMC, y un segundo método basado datos y redes neuronales artificiales.
- ***Evaluar*** la posibilidad y ***viabilidad de implementar este método*** basado en Redes Neuronales ***en la etapa de reconstrucción rápida del HLT.***

LHC y ATLAS

- 2 haces de protones circulan en direcciones opuestas
- La energía de centro de masa en la colisión pp es de $s = 13 \text{ TeV}$



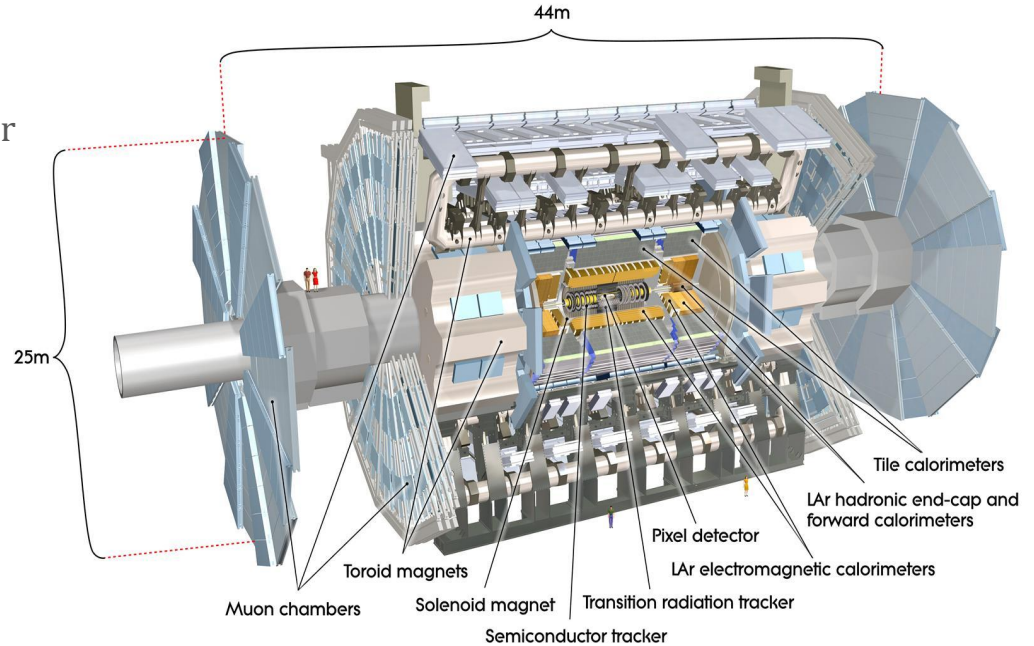
- Los haces se conforman de paquetes de $\sim 10^{11}$ protones c/u.
- Viajan en trenes de hasta 2808 paquetes por tren.
- Tasa de colisiones por paquete de 40 MHz.

El detector ATLAS

Detector multipropósito diseñado para observar partículas producidas en colisiones de protón-protón (pp) e iones-pesados (HI) a altas energías.

Compuesto por los subdetectores:

- Detector Interno de trazas (ID)
- Fino solenoide superconductor
- Calorímetros electromagnéticos (EMC) y hadrónicos (HCal);
- además de un espectrómetro de muones.



Esquema del detector ATLAS donde se pueden observar los distintos subdetectores y regiones que lo conforman.

El detector ATLAS

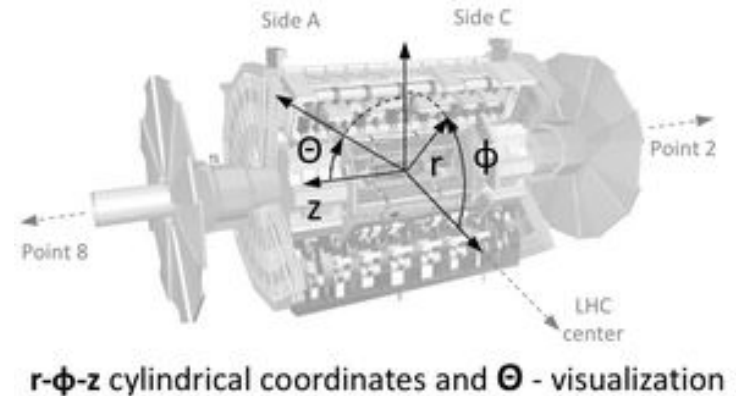
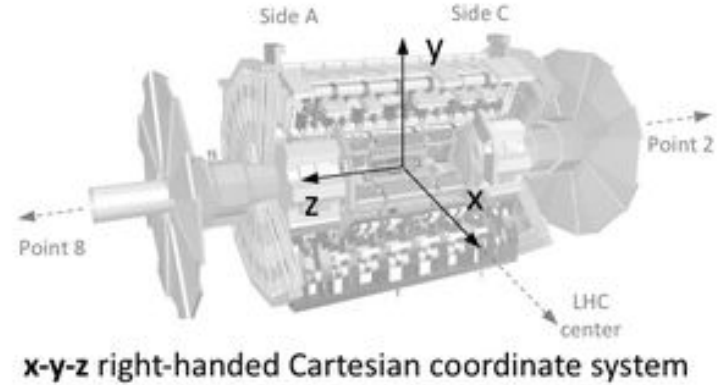
Sistema de coordenadas acorde a la regla de la mano derecha:
cuyo origen es el punto de interacción nominal (*nominal Interaction Point*), que se encuentra en el centro del detector.

$$\eta = -\ln(\tan(\theta/2)),$$

$$\Delta R^2 = (\Delta\eta)^2 + (\Delta\phi)^2.$$

Dos regiones:

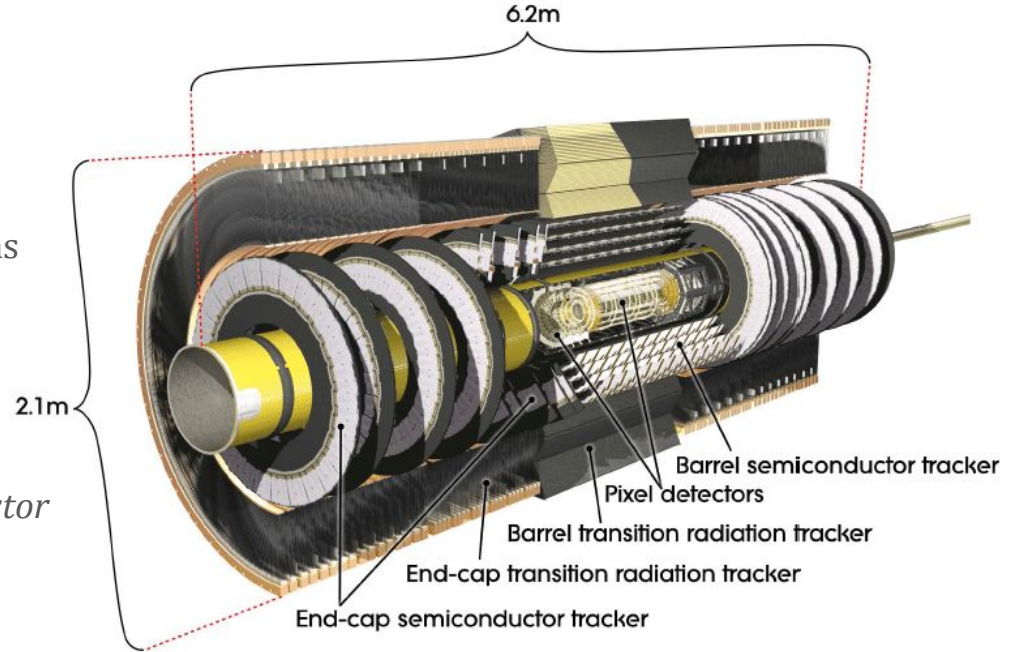
- Barrel: $0 < |\eta| < 1,375$
- Región de transición alrededor de $|\eta| \sim 1,4$.
- Endcap: $1,52 < |\eta| < 3,2$



El detector ATLAS

Detector de Trazas (*Inner tracking Detector*)

- provee la reconstrucción precisa de las trazas de las partículas cargadas, en el rango de pseudorapidez $|\eta| < 2,5$
- Integrado por 4 subdetectores:
IBL, Detector de Píxeles, Detector Semiconductor de Trazas y TRT
- **trazas**: secuencia de puntos coordinados que describen el rastro dejado por partículas cargadas tras interactuar con el detector.

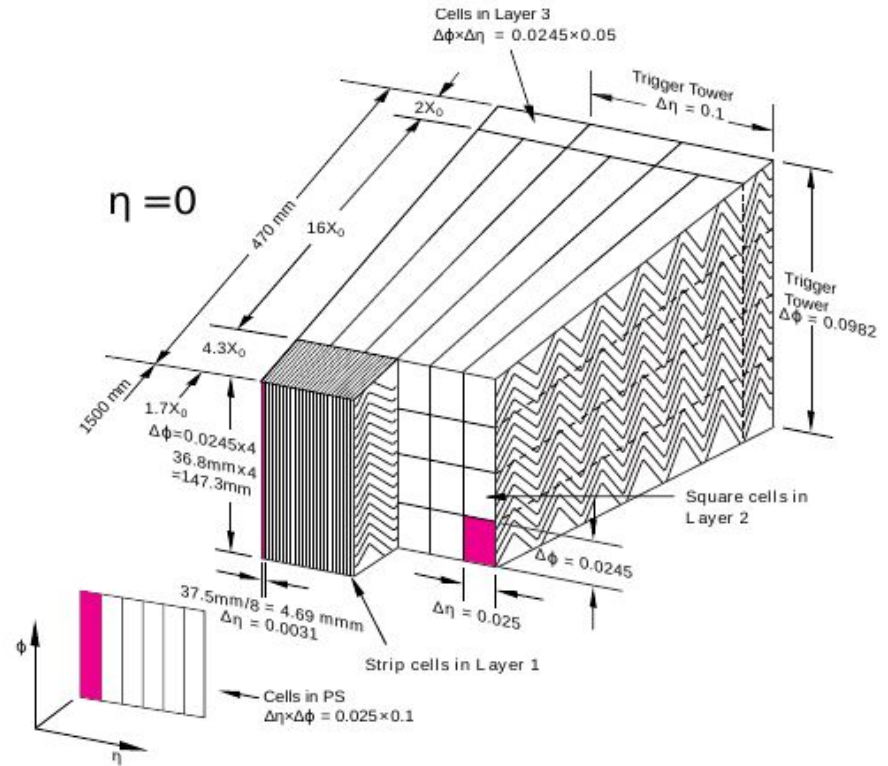


Esquema global del Detector Interno de ATLAS, dónde se puede observar sus componentes y dimensiones.

El detector ATLAS

Calorímetros

- El sistema del calorímetro consta tanto de componentes electromagnéticas (*ElectroMagnetic Calorimeter*) como hadrónicas (*Hadronic CALorimeter*).
- Aporta cobertura total en φ y cobertura de pseudorapidez en el rango $|\eta| < 4,9$; con granularidad más fina en la región asociada al detector interno.

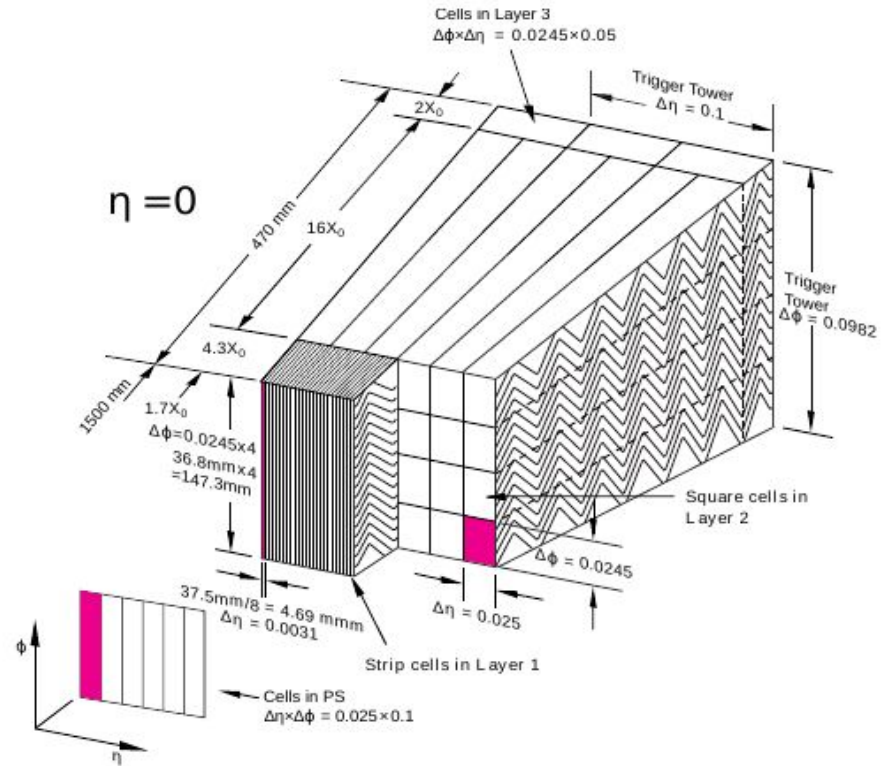


Esquema de la segmentación lateral y longitudinal del calorímetro electromagnético de ATLAS alrededor de $\eta = 0$.

El detector ATLAS

Calorímetro Electromagnético

- Segmentado en 3 capas longitudinales con respecto a la profundidad de las lluvias.
- Dividido en dos medio-tubos o *half-barrels* cubriendo la región de pseudorapidez ($0 < |\eta| < 1,475$), y dos tapas *endcap* cubriendo la región ($1,375 < |\eta| < 3,2$)



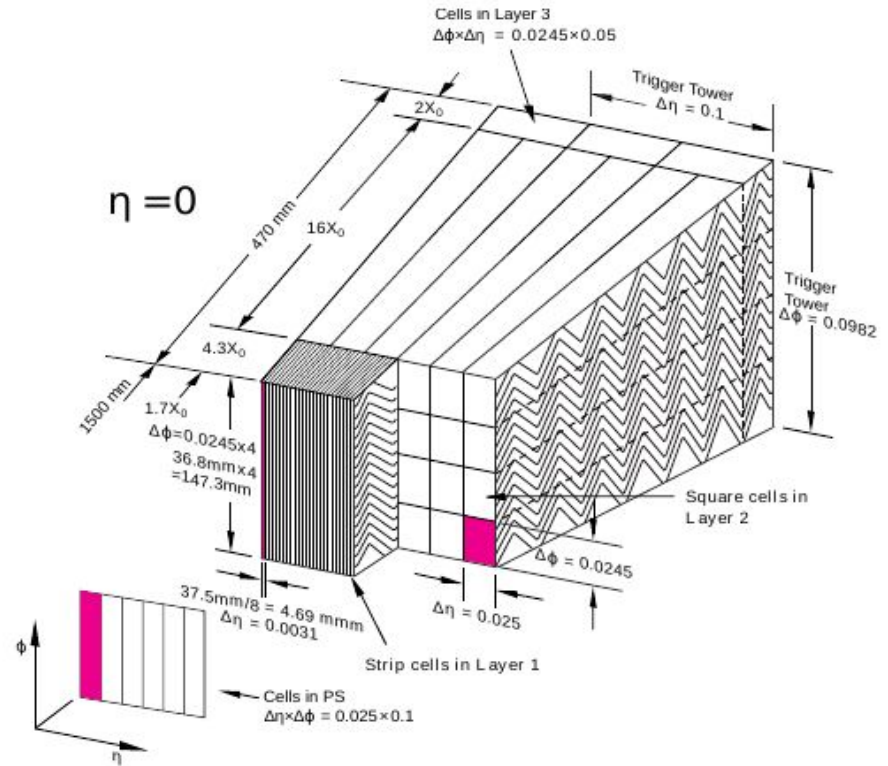
Esquema de la segmentación lateral y longitudinal del calorímetro electromagnético de ATLAS alrededor de $\eta = 0$.

El detector ATLAS

Calorímetro Electromagnético

Capa 2:

- recolecta la mayoría de la energía depositada por electrones y fotones con alta energía transversa.
- Posee un ancho de $\sim 16 X_0$, contando con una granularidad de $0,025 \times 0,0245$ en el espacio ($\Delta\eta, \Delta\phi$)



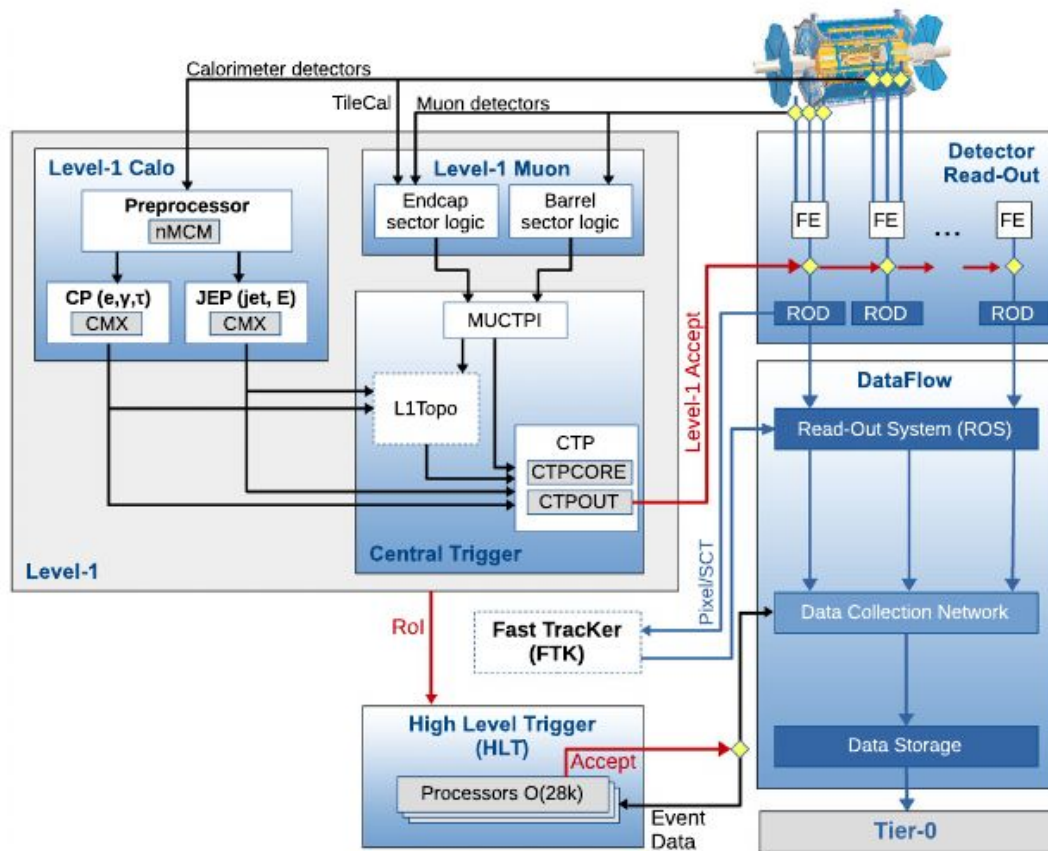
Esquema de la segmentación lateral y longitudinal del calorímetro electromagnético de ATLAS alrededor de $\eta = 0$.

Sistema de disparo (Trigger) de adquisición de datos

Sistema de Trigger

L1

- Implementado en Hardware
- Utiliza señales de gruesa granularidad provenientes de los calorímetros y del espectrómetro de muones.
- Reducir la tasa de eventos de 40 MHz a 100 kHz

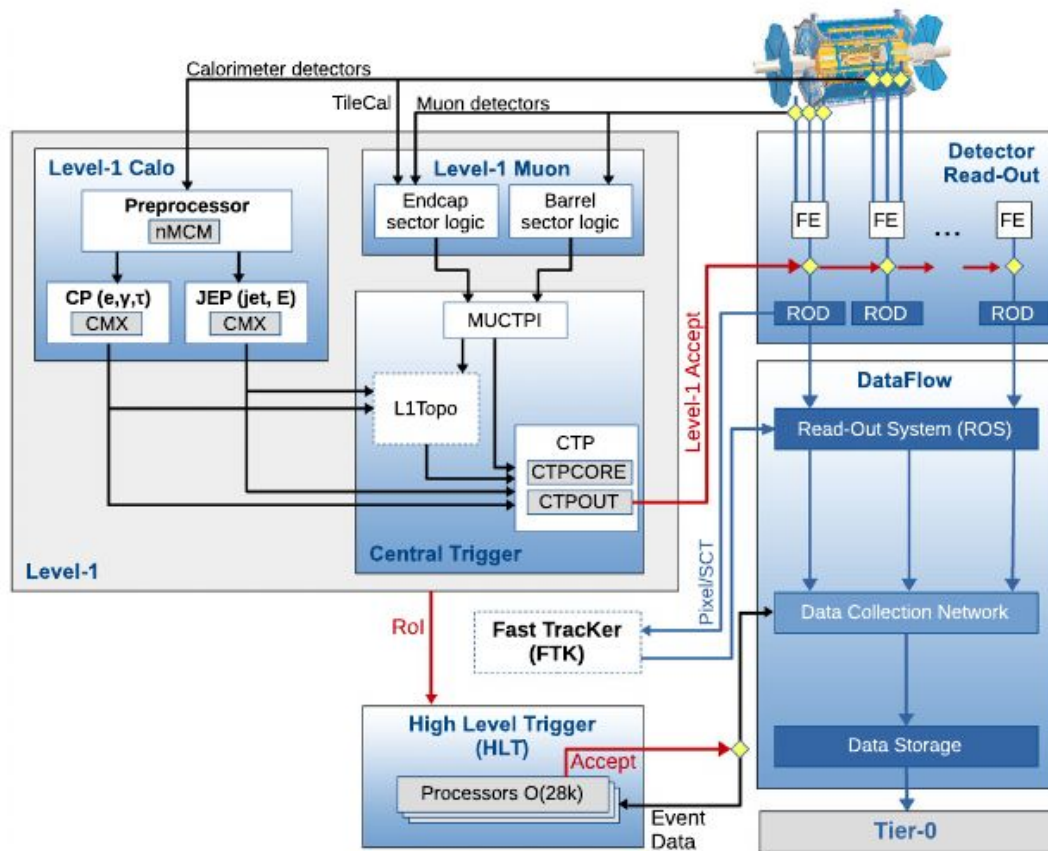


Esquema del sistema TDAQ de ATLAS para Run2.
Las flechas indican la dirección de los datos partiendo del detector hacia las distintas secciones del sistema.

Sistema de Trigger

L1

- Identifican **RoIs** como regiones de 0.2×0.2 en $(\Delta\eta, \Delta\varphi)$
- Los RoIs son zonas del detector donde se obtuvo señal.
- Contienen información de coordenadas y el tipo de candidato (tau, EM, jet).

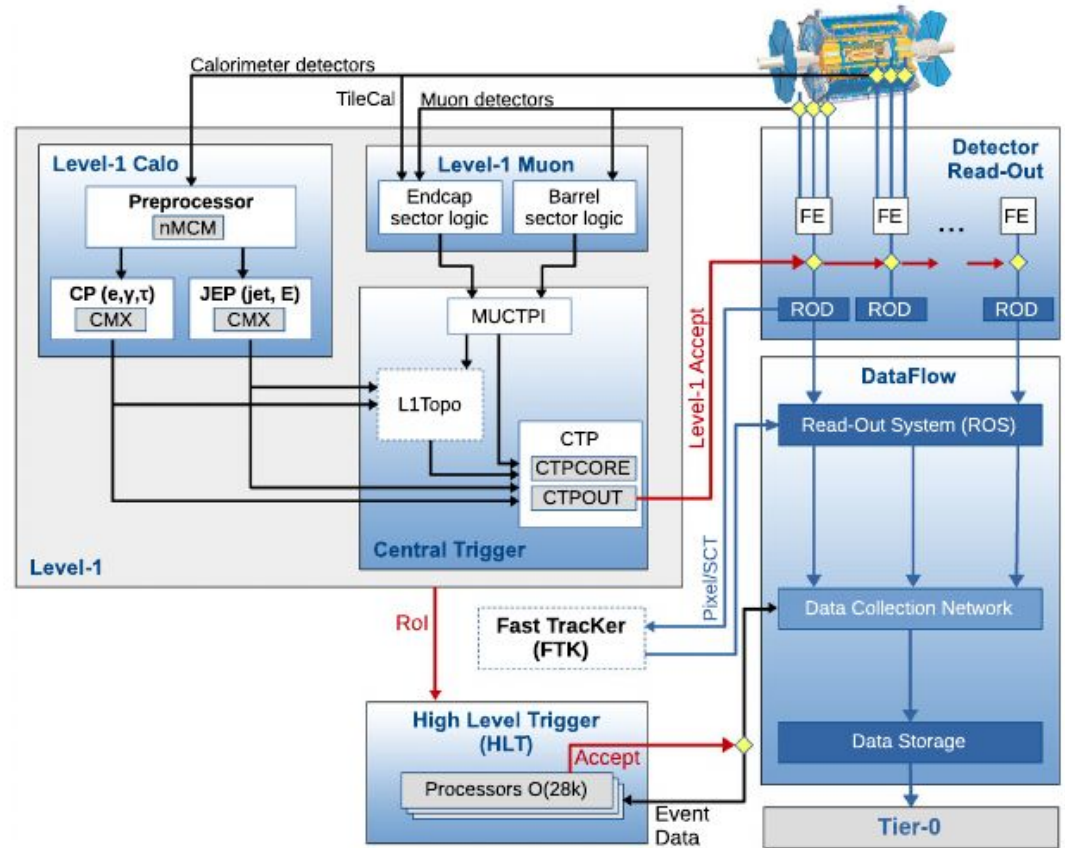


Esquema del sistema TDAQ de ATLAS para Run2.
Las flechas indican la dirección de los datos partiendo del detector hacia las distintas secciones del sistema.

Sistema de Trigger

HLT

- Eventos aceptados por L1 son procesados por el HLT utilizando:
 - información calorimétrica de fina granularidad,
 - medidas de precisión de los espectrómetros de muones,
 - e información de trazas proveniente de ID.

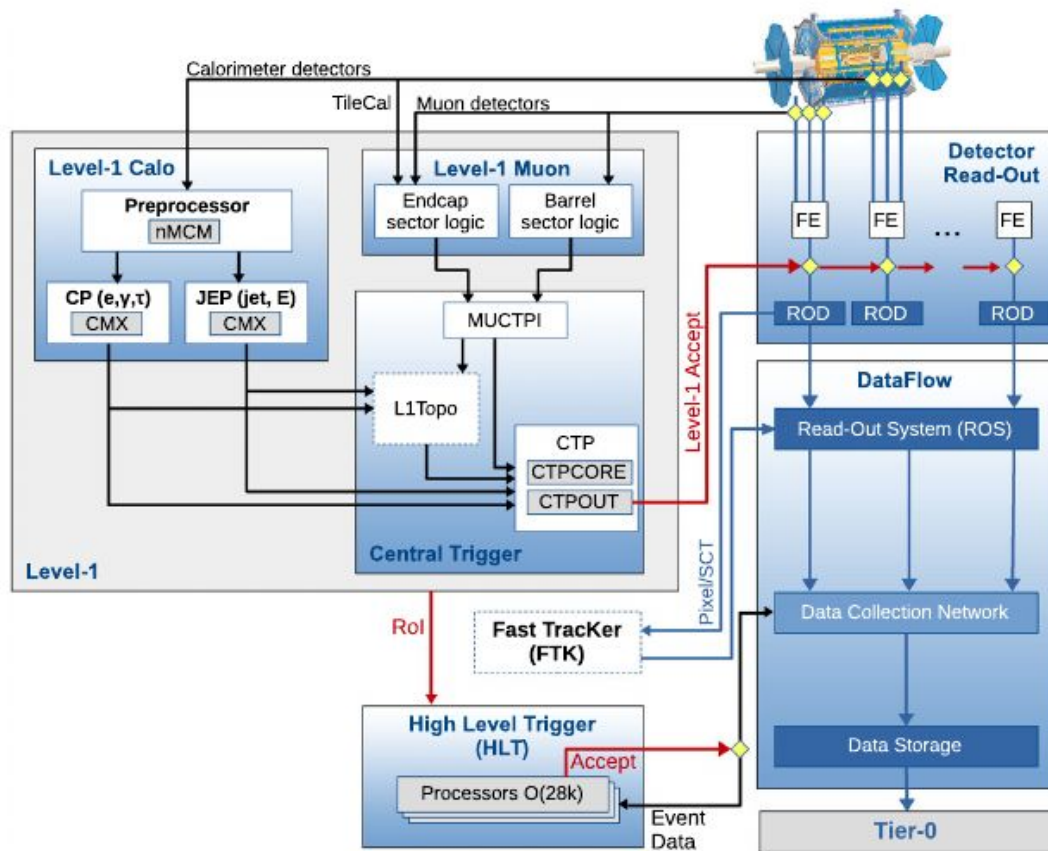


Esquema del sistema TDAQ de ATLAS para Run2.
Las flechas indican la dirección de los datos partiendo del detector hacia las distintas secciones del sistema.

Sistema de Trigger

HLT

- **2 ETAPAS:**
 - Reconstrucción rápida: para hacer el rechazo del fondo lo antes posible.
 - Reconstrucción precisa: similar a *offline*
- Como **resultado** se cuenta con **candidatos partícula** (electrón, fotón, muón, jet,..)

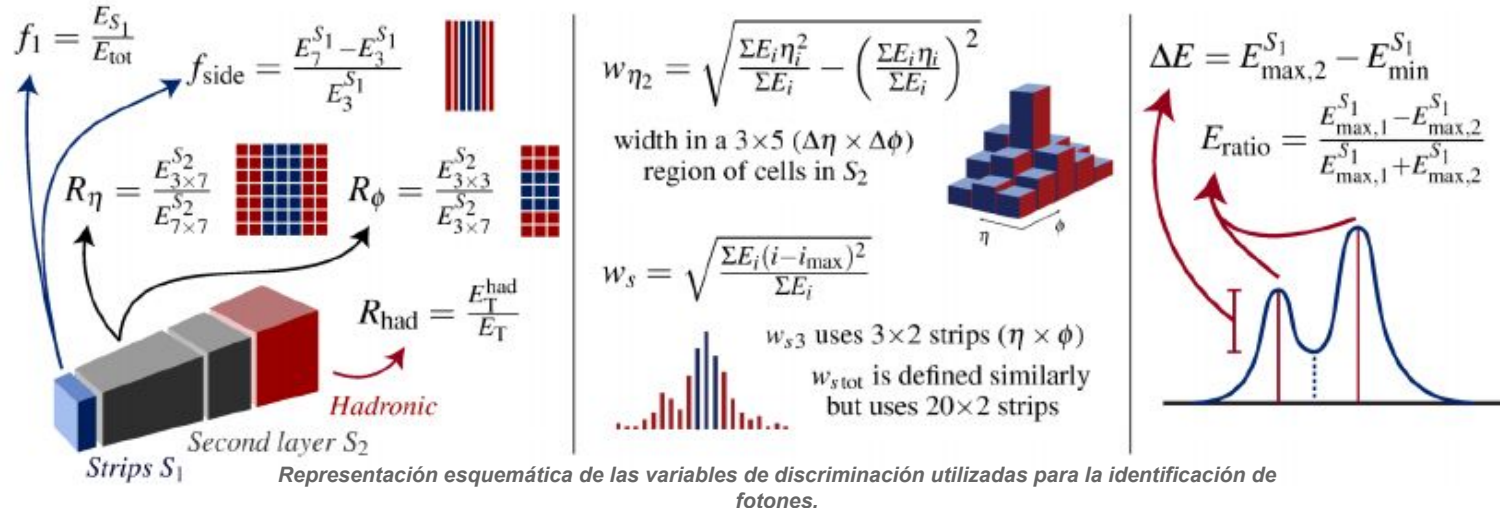


Esquema del sistema TDAQ de ATLAS para Run2.
Las flechas indican la dirección de los datos partiendo del detector hacia las distintas secciones del sistema.

Fotones en ATLAS

Variables de las lluvias electromagnéticas

Describen la forma y las propiedades asociadas a lluvias electromagnéticas (**shower shapes**) depositadas en los calorímetros, las cuales proveen buena separación entre fotones prompt y fotones falsos.



Reconstrucción

offline de electrones y fotones

- utiliza **topo-clusters**: clusters dinámicos de tamaño variable de los depósitos de energía medidos en celdas del EMC y HCAL topológicamente conectadas.
- correcciones a la posición inicial y calibración de energía a los topo-clusters, son asignados a trazas de ID.
- *topo-clusters* que no fueron asignados a ninguna traza o vértice de conversión son reconstruidos como candidatos-fotones.

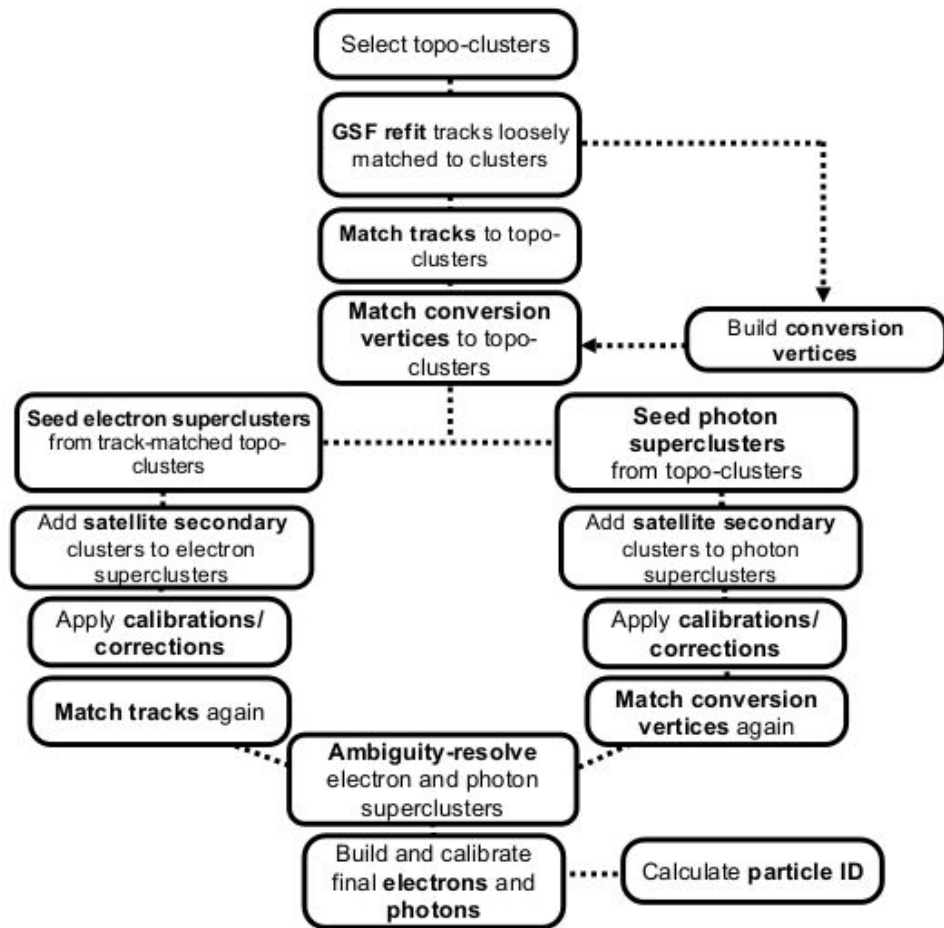


Diagrama de flujo del proceso de reconstrucción offline de objetos en ATLAS.

Reconstrucción

online

- en el HLT no se utiliza la información del detector de trazas para identificar conversiones.
- en el trigger de ATLAS la reconstrucción de fotones se hace solamente con la información de los calorímetros bajo la hipótesis de fotones no convertidos.

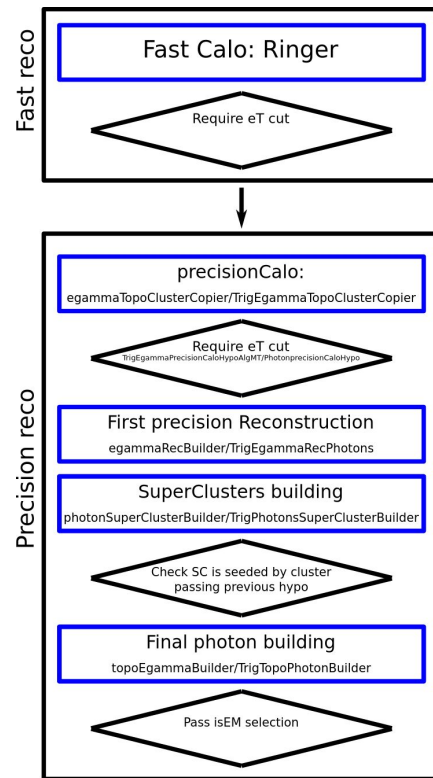
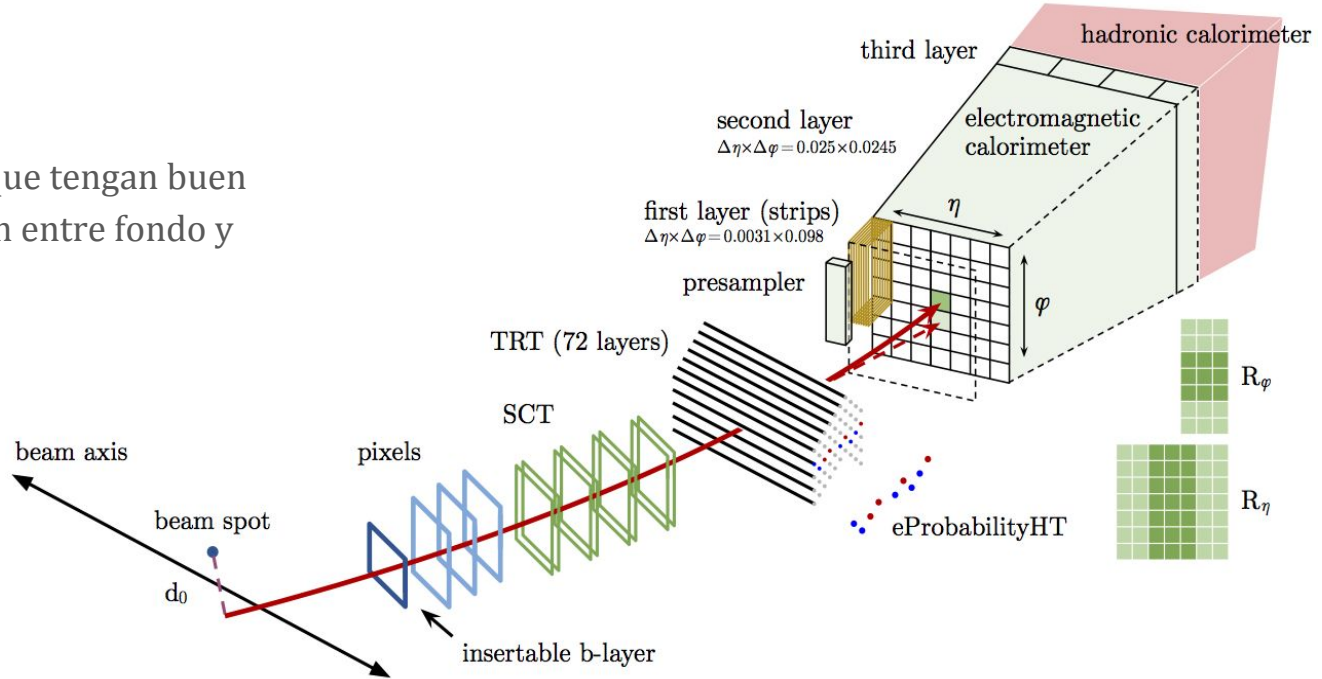


Diagrama de flujo del proceso de reconstrucción online de fotones en ATLAS.

Reconstrucción rápida en el HLT

online

- Se aplican selecciones que tengan buen poder de discriminación entre fondo y señal:
 - E_T del cluster
 - shower shapes.
 E_{ratio} , R_{had} y R_{eta}

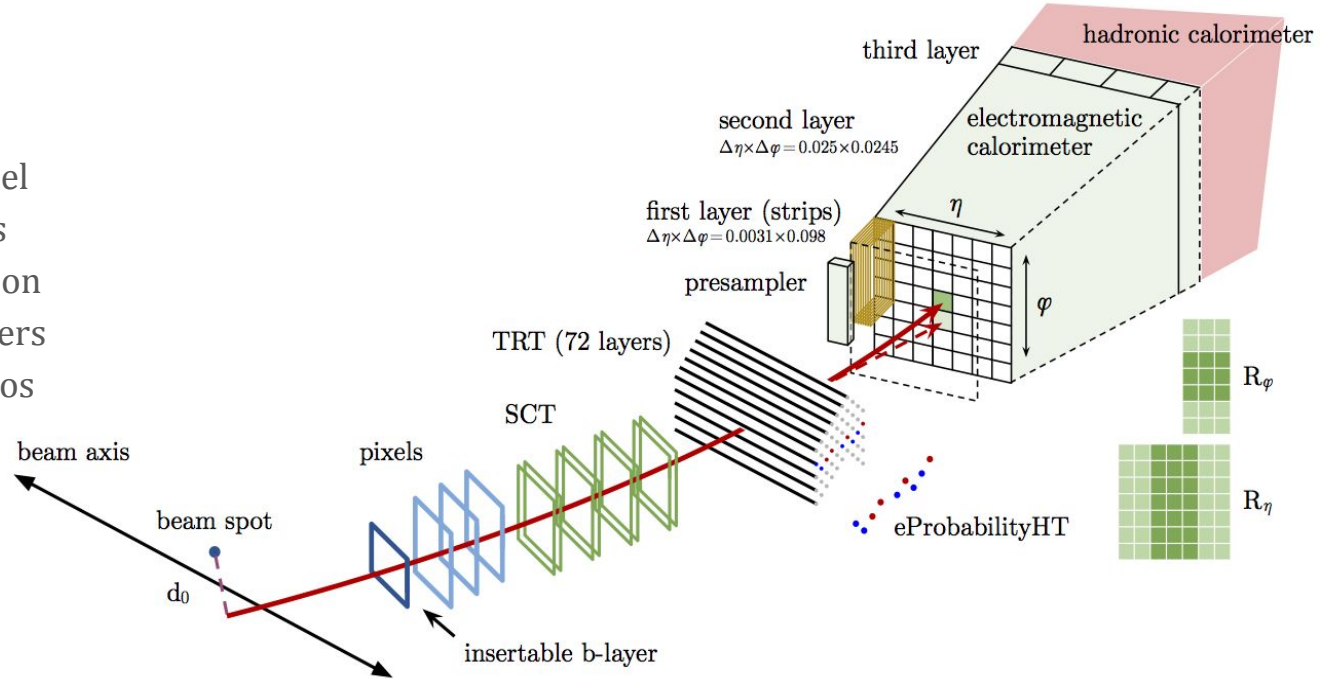


Esquema del proceso de reconstrucción de los objetos a través del detector

Reconstrucción de precisión en el HLT

online

- Los clusters de celdas del EMC que dan lugar a los candidatos fotones no son *topo-clusters* si no clusters rectangulares producidos por un algoritmo de *sliding-window*.



Esquema del proceso de reconstrucción de los objetos a través del detector

Identificación de fotones

Offline y online

- Depende de:
 - requisitos de selección rectangulares** online y offline, *basado en variables de ShowerShapes*.
 - criterios de aislamiento** de los candidatos a otras partículas del evento.
 - Las selecciones se optimizan por separado para candidatos fotones convertidos y no convertidos.
- Para fotones se definen tres puntos de trabajo de identificación (*identification Working Points*): “loose”, “medium”(HLT only) y “tight”.

Category	Description	Name	loose	tight
Acceptance	$ \eta < 2.37$, with $1.37 \leq \eta < 1.52$ excluded	—	✓	✓
Hadronic leakage	Ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ or $ \eta > 1.52$)	R_{had_1}	✓	✓
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	✓	✓
EM middle layer	Ratio of the energy in $3 \times 7 \eta \times \phi$ cells over the energy in 7×7 cells centered around the photon cluster position	R_η	✓	✓
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	w_{η_2}	✓	✓
EM strip layer	Ratio of the energy in $3 \times 3 \eta \times \phi$ cells over the energy of 3×7 cells centered around the photon cluster position	R_ϕ		✓
	Lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2)/(\sum E_i)}$, where i runs over all strips in a window of $3 \times 2 \eta \times \phi$ strips, and i_{max} is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit	$w_{s\ 3}$		✓
	Total lateral shower width $\sqrt{(\sum E_i (i - i_{\text{max}})^2)/(\sum E_i)}$, where i runs over all strips in a window of $20 \times 2 \eta \times \phi$ strips, and i_{max} is the index of the highest-energy strip measured in the strip layer	$w_{s\ \text{tot}}$		✓
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips	f_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima	ΔE_s		✓
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies	E_{ratio}		✓
	Ratio of the energy in the first layer to the to the total energy of the EM cluster	f_1		✓

Tabla con la definición de algunas de las variables que describen la distribución de la lluvias.

Identificación de fotones

Offline y online

- Caracterizan la evolución de los rastros laterales y longitudinales en el EMC, como también la fracción de lluvia perdida en HCAL.

Identificación de electrones

Offline y online

- Método de likelihood para electrones. Se definen tres WP:
LHloose, *LHmedium* y *LHtight*

Category	Description	Name	loose	tight
Acceptance	$ \eta < 2.37$, with $1.37 \leq \eta < 1.52$ excluded	—	✓	✓
Hadronic leakage	Ratio of E_T in the first sampling layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ or $ \eta > 1.52$)	R_{had1}	✓	✓
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	✓	✓
EM middle layer	Ratio of the energy in $3 \times 7 \eta \times \phi$ cells over the energy in 7×7 cells centered around the photon cluster position	R_η	✓	✓
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	w_{η_2}	✓	✓
EM strip layer	Ratio of the energy in $3 \times 3 \eta \times \phi$ cells over the energy of 3×7 cells centered around the photon cluster position	R_ϕ		✓
	Lateral shower width, $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $3 \times 2 \eta \times \phi$ strips, and i_{max} is the index of the highest-energy strip calculated from three strips around the strip with maximum energy deposit	w_{s3}		✓
	Total lateral shower width $\sqrt{(\sum E_i (i - i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $20 \times 2 \eta \times \phi$ strips, and i_{max} is the index of the highest-energy strip measured in the strip layer	w_{stot}		✓
	Energy outside the core of the three central strips but within seven strips divided by energy within the three central strips	f_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer and the energy reconstructed in the strip with the minimum value found between the first and second maxima	ΔE_s		✓
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in the secondary maximum in the cluster to the sum of these energies	E_{ratio}		✓
	Ratio of the energy in the first layer to the to the total energy of the EM cluster	f_1		✓

Tabla con la definición de algunas de las variables que describen la distribución de la lluvias.

Aislamiento

a candidatos fotones y leptones

- *Propósito: **reducir el fondo*** de jets mal identificados y de decaimiento de hadrones livianos y pesados dentro de los jets.
- Se define el aislamiento de fotones (photon isolation) basado en la energía transversa de un cono centrado alrededor de la dirección del candidato-fotón de tamaño angular ΔR . Está caracterizada por 2 cantidades, la aislación calorimétrica E_T^{iso} (calorimeter isolation) y la aislación de trazas p_T^{iso} (track isolation).
- La aislación E_T^{iso} se calcula como la **suma de energías transversas de los topo-clusters en los calorímetros en un cono alrededor del candidato**; luego de sustraer evento por evento la energía depositada por los candidatos-fotones, la contribución de los eventos subyacentes y el pile-up.

Redes Neuronales Artificiales

Redes Neuronales Artificiales

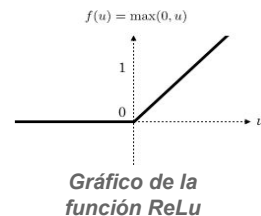
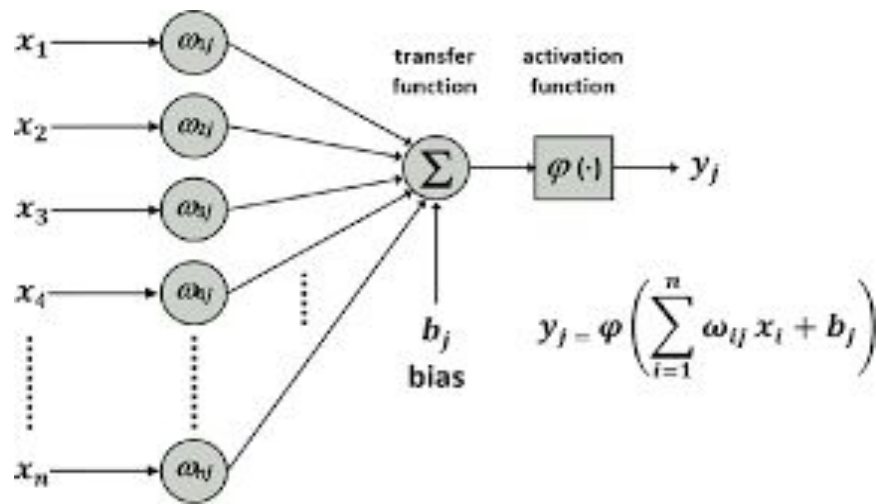
El término Red Neuronal Artificial (Artificial Neural Network) hace referencia a un algoritmo computacional basado en neuronas artificiales, el nodo elemental del algoritmo. Las cuales trabajan entrelazados de manera distribuida para aprender colectivamente a partir de los datos de entrada con el objetivo de optimizar el resultado de salida.

x: vector de entrada

y: vector de salida

w: matriz de pesos

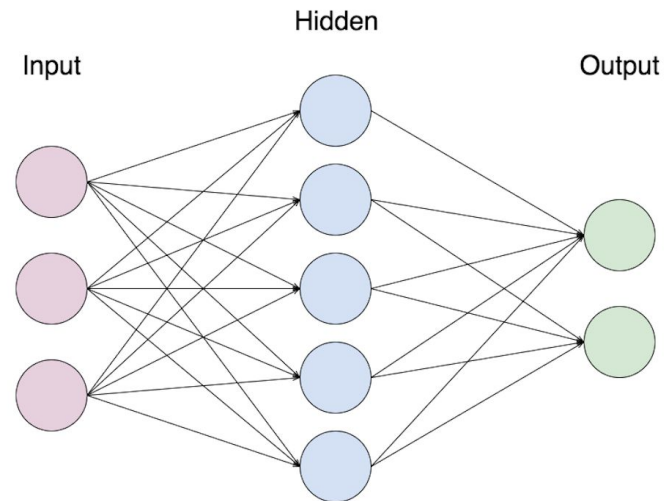
b: sesgo



Esquema de una neurona artificial.

Redes Neuronales Artificiales

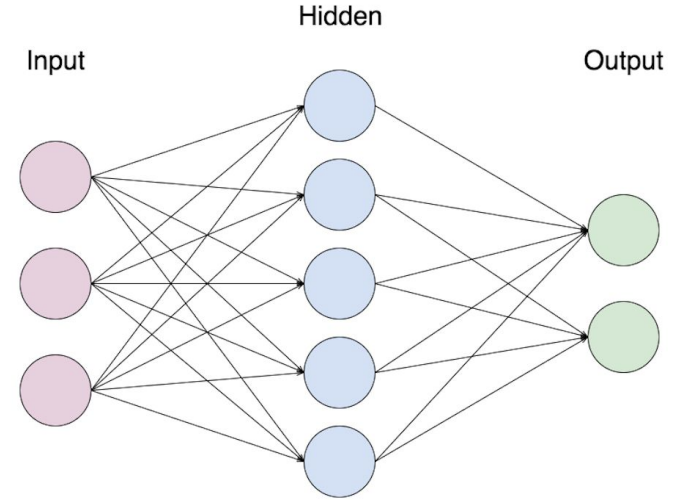
- **capas:** sucesivas conexiones de nodos.
- **modelo:** un algoritmo con dos o más capas conectadas.
- El concepto de **aprendizaje** hacer referencia al labor que realizan las capas internas de la red, las cuales toman decisiones de capas anteriores y pesan cómo un cambio estocástico mejora o empeora el resultado final.



Esquema de una red neuronal artificial, compuesta por tres capas.

Redes Neuronales Artificiales

- 2 paradigmas de aprendizaje: Aprendizaje *Supervisado* y Aprendizaje *No-supervisado*.
- El *Aprendizaje Supervisado* se caracteriza por requerir de datos de entrada etiquetados que actúen como “objetivos” para el entrenamiento de la red.



Esquema de una red neuronal artificial, compuesta por tres capas.

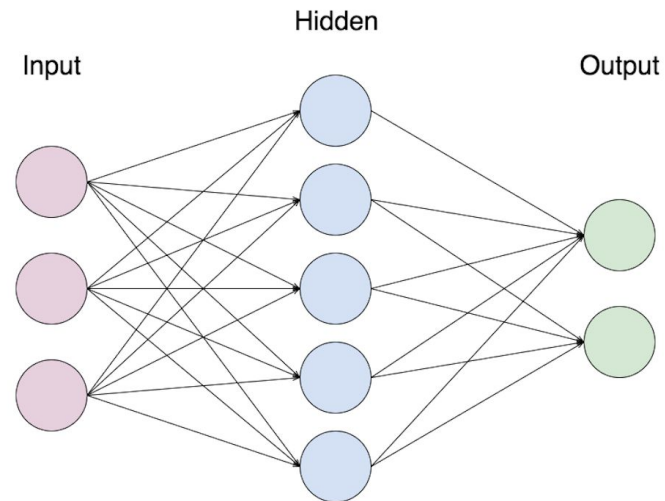
Redes Neuronales Artificiales

Proceso de entrenamiento:

1. Inicializar la matriz de pesos aleatoriamente.
2. Procesar la señal desde la capa de entrada hasta obtener un resultado final en la capa de respuesta.
3. Evaluar la función de pérdida entre los resultados del 2^{do} paso frente al resultado esperado.

La función de pérdida calcula qué tan apartado del valor esperado es la predicción realizada por la red

$$\epsilon(\mathbf{w}) = - \sum_{n=1}^N t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n)$$



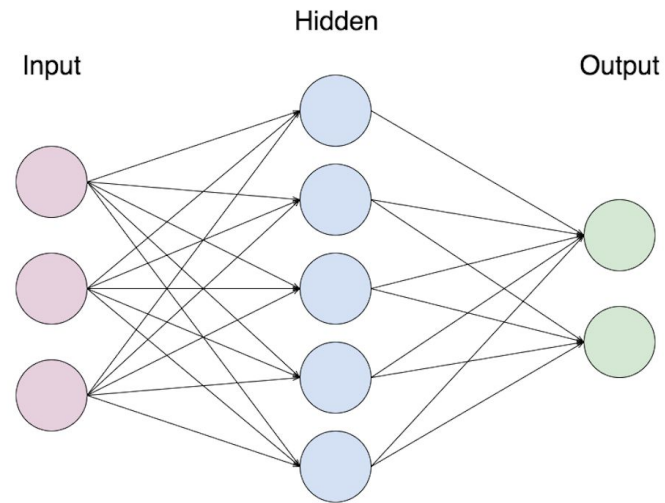
Esquema de una red neuronal artificial, compuesta por tres capas.

Redes Neuronales Artificiales

4. Propagación inversa (Backpropagation): aplicar un reajuste de la matriz de pesos en el sentido negativo de la propagación de la señal que minimice la función de pérdida.

La ejecución de 2^{do} paso al 4^{to} conforma lo que se conoce como una **época** de entrenamiento.

La recursividad del algoritmo implica entrenar durante la cantidad de épocas necesarias hasta lograr minimizar la función de pérdida por debajo de la cota deseada.



Esquema de una red neuronal artificial, compuesta por tres capas.

Redes Neuronales Convolucionales

3 tipos de capas:

- **convolucionales:** constituidas por filtros (*kernels*). Aplica una convolución entre los datos y cada uno de los filtros para producir un mapa de activación 2D (activation map). Luego aplica una función de activación a la salida de la capa previa.

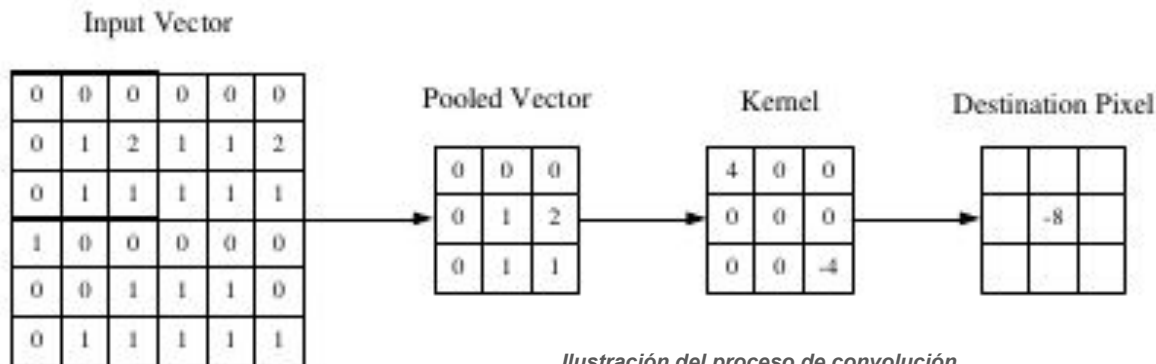


Ilustración del proceso de convolución.

Redes Neuronales Convolucionales

3 tipos de capas:

- **agrupamiento:** opera sobre cada mapa de activación de la entrada rescaleando su dimensión mediante alguna función de agrupamiento, comúnmente máximo o promedio, para reducir el número de parámetros y la complejidad computacional del modelo.

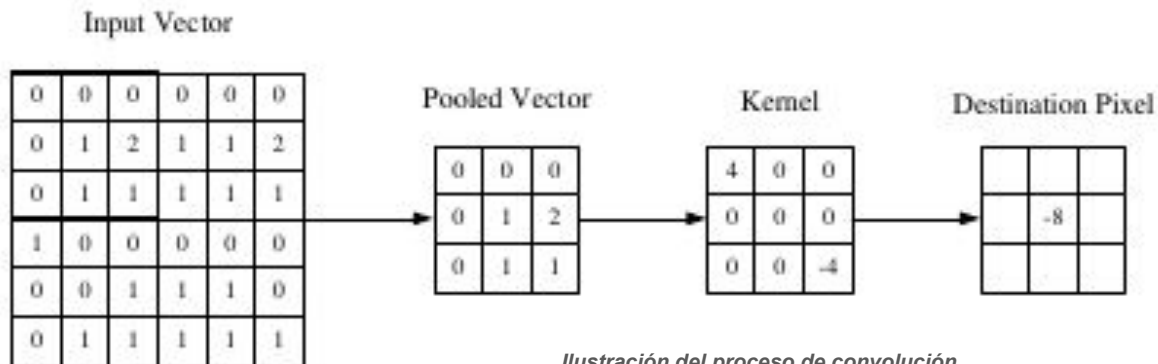
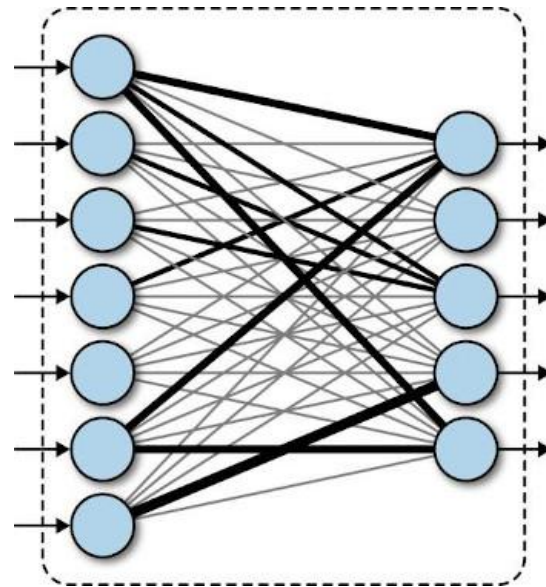


Ilustración del proceso de convolución.

Redes Neuronales Convolucionales

3 tipos de capas:

- ***densas***: Se caracteriza por contener neuronas que están directamente conectadas a las neuronas de las dos capas adyacentes (anterior y posterior), sin estar mutuamente conectadas.



Esquema de una capa densa

Identificación de fotones con CNN

Identificación de fotones con CNN

Idea general del método

- Utilizamos CNN porque realizan una buena tarea en el reconocimiento de patrones de imágenes.
- Definimos dos categorías: señal y fondo.
- Para cada evento nos fabricamos una imagen a partir de las celdas del calorímetro que forman los topo-clusters de los fotones.
- Comparamos los métodos de clasificación: *tight* vs *CNN*.

Identificación de fotones con CNN

Almacenamiento de eventos

- Datos medidos en la región barrel de la señal depositada en la 2^{da} capa del EMC, debido a que uno de los objetivos finales de este estudio es probar la técnica en el FastReco del HLT .
- Desarrollamos un software de análisis implementado como un paquete de ATHENA, llamado [CellDumpingPackage](#).
 - Levanta de los eventos: los candidatos partícula, cantidades asociadas relevantes a estos, aplica la selección Zrad, y obtiene las celdas que conforman los topo-clusters.

Identificación de fotones con CNN

Creando las imágenes de los *topo-clusters*

- Un cluster está formado por las celdas del calorímetro dónde la partícula depositó su energía, es decir que a cada celda la podemos entender como un vector almacenando:
celda = (η, ϕ, E) .
- Dado que las coordenadas $(\eta, \phi) \in \mathbb{R}$ y que los índices de una matriz son \mathbb{N} , es necesario dar una regla de transformación. Tomamos como referencia los anchos de las celdas de la segunda capa del calorímetro, donde establece que $(\Delta\eta, \Delta\phi) = (0.025, 0.0245)$.
- Calculamos cuál es la distancia del cluster al origen de coordenadas $(\eta_0, \phi_0) = (0, 0)$, mediante las siguientes ecuaciones

$$step_{\eta} = round\left(\frac{(\eta + \Delta\eta/2) * sig(\eta)}{\Delta\eta} + \delta_{\eta}\right),$$

$$step_{\phi} = round\left(\frac{(\phi + \Delta\phi/2) * sig(\phi)}{\Delta\phi} + \delta_{\phi}\right).$$

Identificación de fotones con CNN

Creando las imágenes de los *topo-clusters*

- Finalmente se normaliza la distribución de energía de la matriz, con el fin de evitar introducir un sesgo en p_T a la hora de entrenar la red neuronal.
- Imágenes de dimensiones que rondan 10×30 celdas.

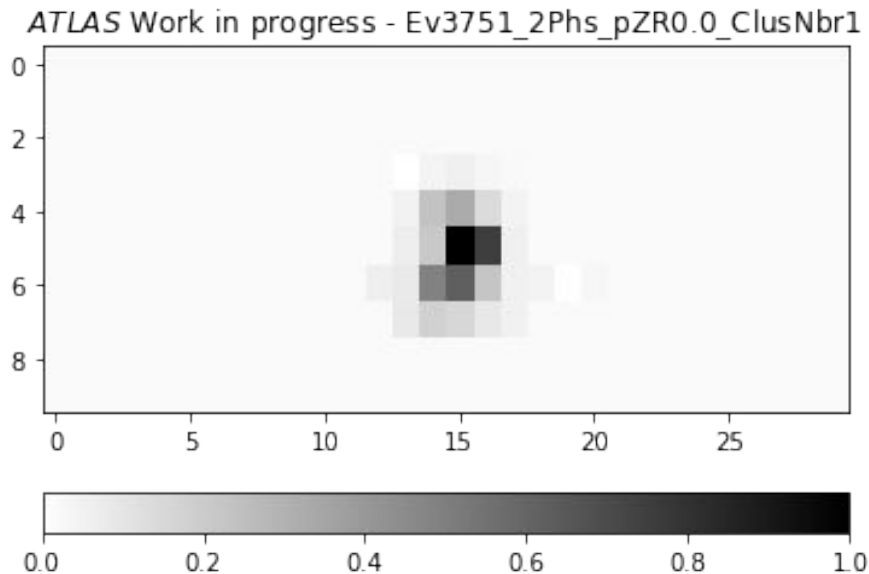
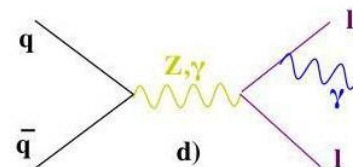
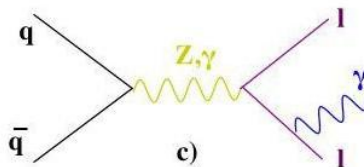
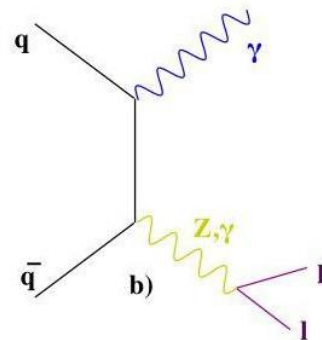
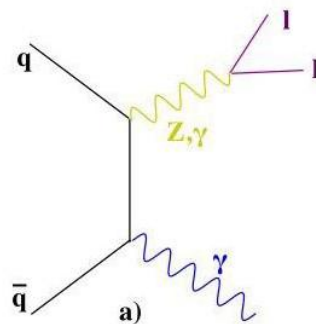


Imagen de un topo-cluster de 10x30 donde cada celda representa la fracción de energía depositada por el candidato fotón respecto a la energía total del cluster.

Identificación de fotones con CNN

Definición y obtención de señal

- Se utilizaron fotones reconstruidos del decaimiento radiativo del bosón Z, (en particular el canal $Z \rightarrow e^+ e^-$).
- Esto se debe a que es posible obtener una muestra de eventos con poco fondo, aplicando tan sólo unos pocos criterios de selección, y por ende contar con una muestra de señal poco contaminada.



Diagramas de Feynman en el SM del decaimiento radiativo del Z. Los diagramas a) y b) con Initial State Radiation. Los diagramas c) y d) son Final State Radiation.

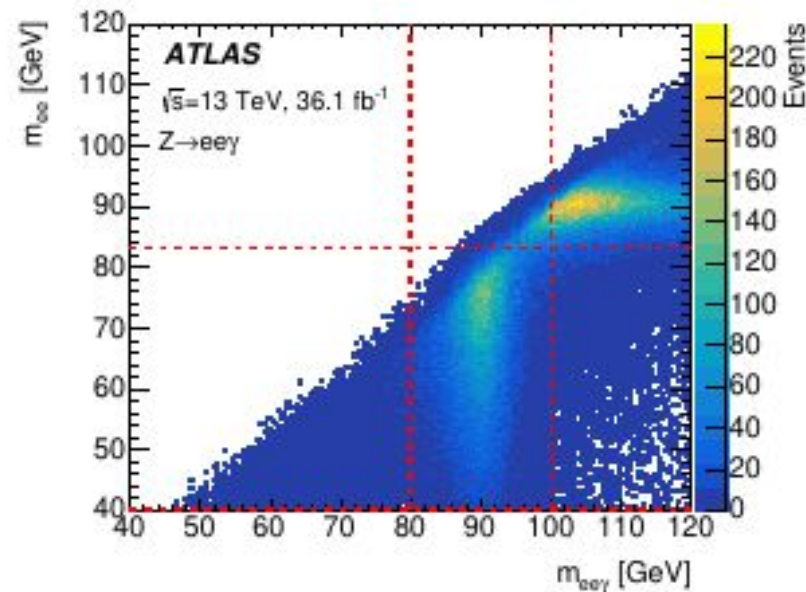
Identificación de fotones con CNN

Decaimiento radiativo del bosón Z

Se utilizaron eventos disparado por triggers de electrones, con el objetivo de contar con fotones no sesgados por el trigger.

Estos triggers requieren:

- 1 candidato electrón (e^\pm) aislado que pase la selección *LHtight* con al menos 26 GeV de momento transversal, o 2 candidatos electrones que pasen la selección *LHloose* con al menos 17 GeV de momento transversal cada uno.



Distribución de $m_{ee\gamma}$ vs m_{ee} para eventos que satisfacen todos los criterios de selección Z→eeγ excepto los de $m_{ee\gamma}$ y m_{ee} .

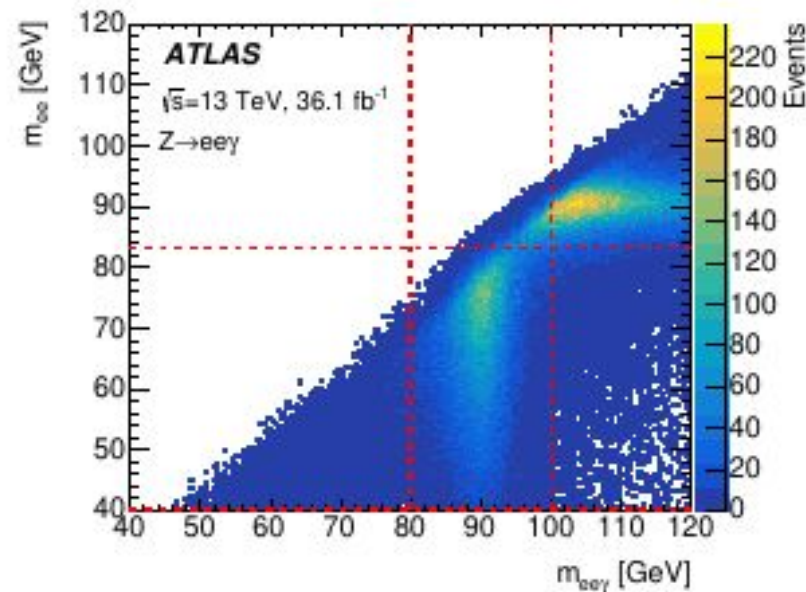
Identificación de fotones con CNN

Decaimiento radiativo del bosón Z

Se utilizaron eventos disparado por triggers de electrones, con el objetivo de contar con fotones no sesgados por el trigger.

Estos triggers requieren:

- En la reconstrucción offline del evento del decaimiento radiativo del bosón Z requiere la presencia de un candidato-fotón (γ) y un par leptónico de carga opuesta de candidatos electrones (e^\pm).



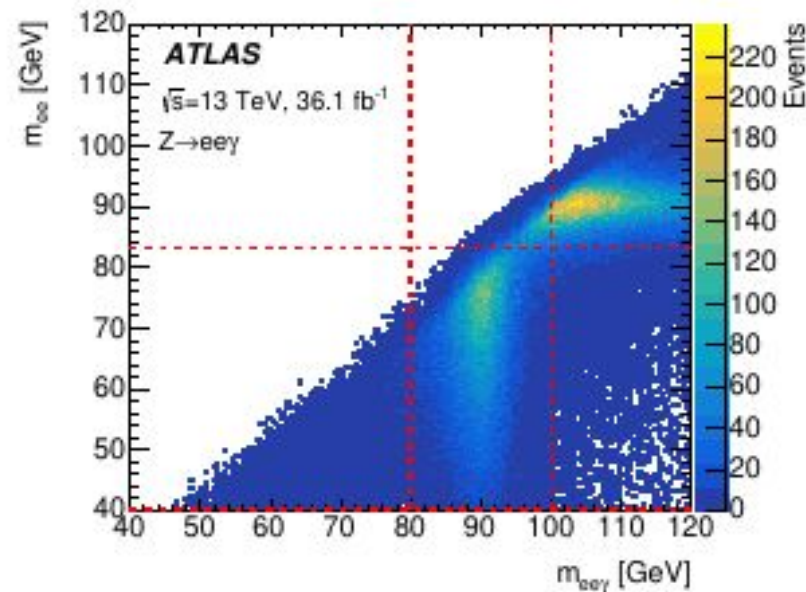
Distribución de $m_{ee\gamma}$ vs m_{ee} para eventos que satisfacen todos los criterios de selección Z \rightarrow e e γ excepto los de $m_{ee\gamma}$ y m_{ee} .

Identificación de fotones con CNN

Decaimiento radiativo del bosón Z

los requisitos que deben satisfacer el *sistema dileptónico* y el *sistema de los tres objetos* son:

1. $40 \text{ GeV} < m_{\ell\ell} < 83 \text{ GeV}$
2. $80 \text{ GeV} < m_{\ell\ell\gamma} < 100 \text{ GeV}$,
3. $\Delta R > 0,4$, entre el fotón y el electrón más cercano.



Distribución de $m_{ee\gamma}$ vs m_{ee} para eventos que satisfacen todos los criterios de selección $Z \rightarrow ee\gamma$ excepto los de $m_{ee\gamma}$ y m_{ee} .

Identificación de fotones con CNN

Decaimiento radiativo del bosón Z

los requisitos que deben satisfacer los *candidatos electrones* son:

1. $E_T^{e^\pm} > 10 \text{ GeV}$,
2. $|\eta^{e^\pm}| < 1,37$ ó $1,52 \leq |\eta^{e^\pm}| < 2,37$.
3. *LHmedium*
4. *aislación de trazas loose*

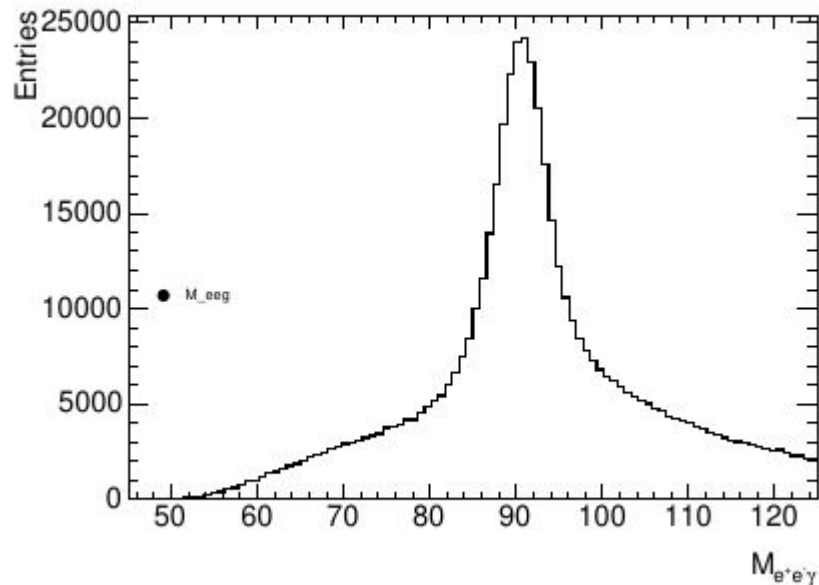


Gráfico de la masa invariante del fotón y el par electrón positrón, cumpliendo la selección Zrad salvo por los cortes en masa para $M_{ee\gamma}$

Identificación de fotones con CNN

Decaimiento radiativo del bosón Z

los requisitos que deben satisfacer los *candidatos fotones* son:

1. $E_T^\gamma > 10 \text{ GeV}$,
2. $|\eta^\gamma| < 1,37$
3. cumplir *loose*

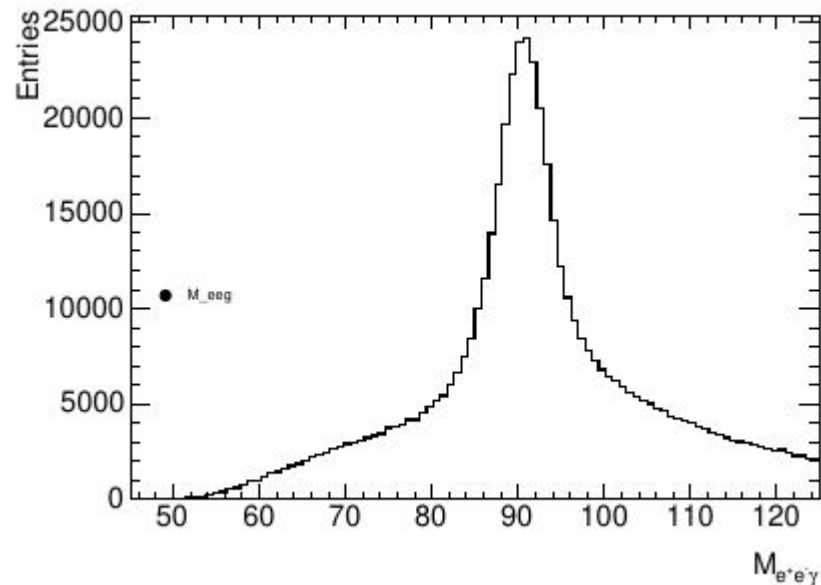
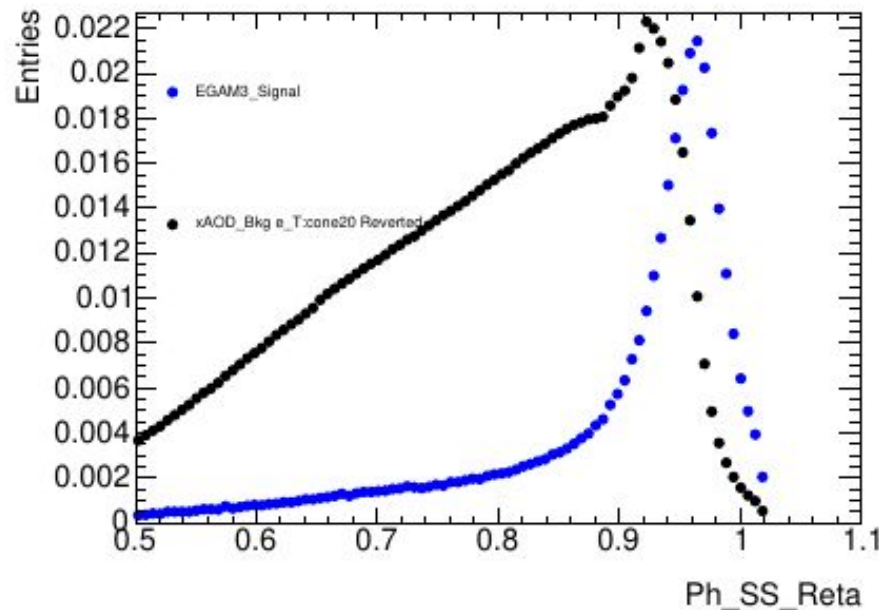


Gráfico de la masa invariante del fotón y el par electrón positrón, cumpliendo la selección Zrad salvo por los cortes en masa para $M_{ee\gamma}$

Identificación de fotones con CNN

Definición y obtención de fondo

- El fondo más importante está dado por jets de partículas en los que un π^0 contiene la mayoría de la energía del jet.
- El decaimiento del π^0 da un par de fotones muy cercanos que pueden dejar en el detector una señal muy parecida a la de un fotón real



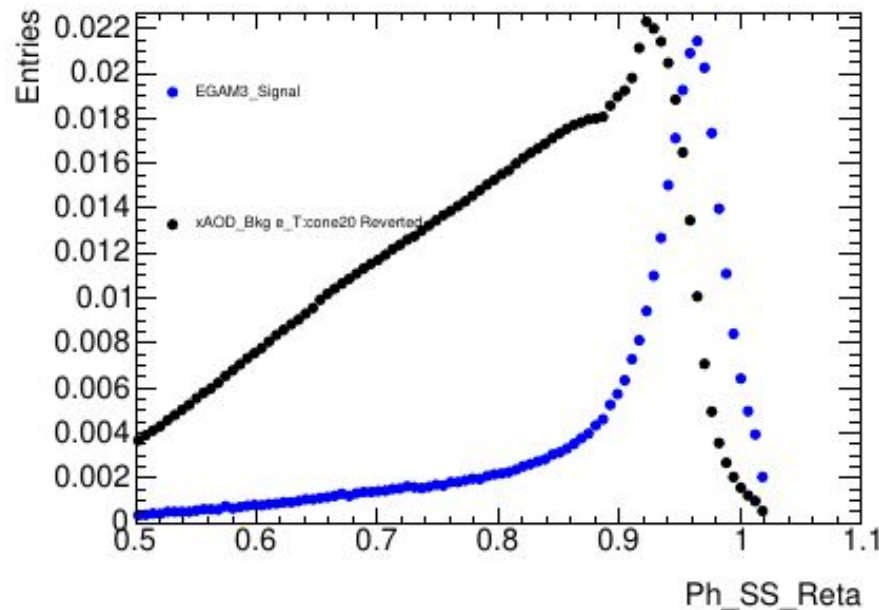
Histogramas normalizados en función de la variable R_n para eventos clasificados como señal (azul) y fondo (negro), mediante aislamiento invertido E_T^{cone20}

Identificación de fotones con CNN

Definición y obtención de fondo

- *Sí se requiere que los candidatos fotones no estén asilados se consigue una muestra de falsos fotones a partir de datos.*
- Por lo tanto, para obtener una muestra enriquecida en fondo se utilizan criterios de aislamiento calorimétrico invertido

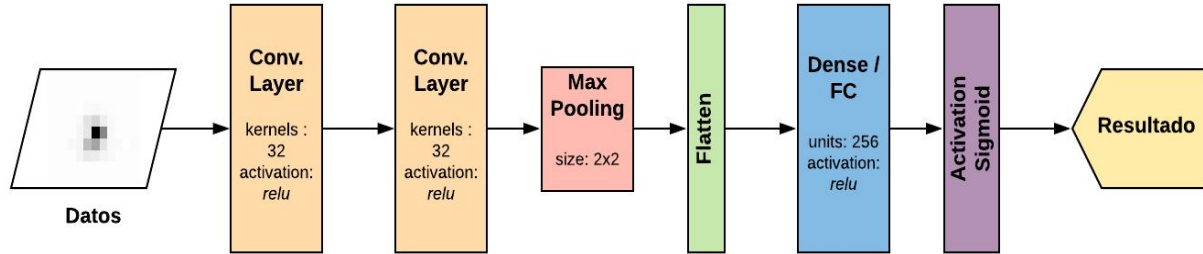
$$E_T^{\text{cone20},\gamma} > p_T^\gamma \times 0,07,$$



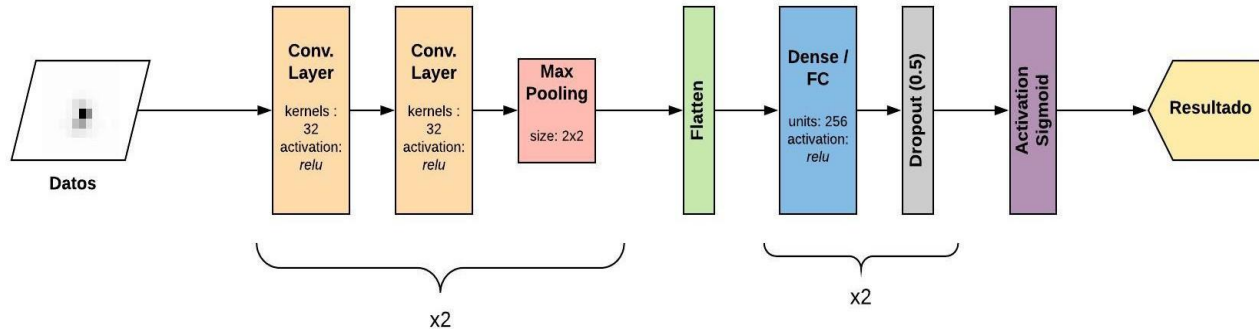
Histogramas normalizados en función de la variable R_n para eventos clasificados como señal (azul) y fondo (negro), mediante aislamiento invertido E_T^{cone20}

Identificación de fotones con CNN

Modelo 0



Modelo 1



Esquema de la arquitectura de CNN utilizada para los dos modelos entrenados.
Arriba el modelo 0, abajo el modelo 1.

Identificación de fotones con CNN

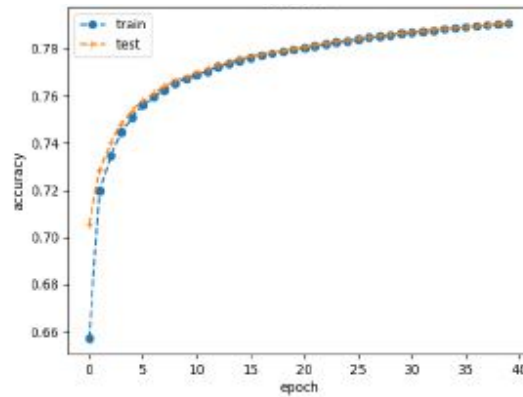
Entrenando las CNNs

- Conjunto de entrenamiento: 30000 imágenes de señal y 30000 de fondo.
- métrica: *precisión*
- optimizador: *RMSprop* ($lr = 0.0001$)
- función de pérdida: *entropía cruzada*
- épocas: 40

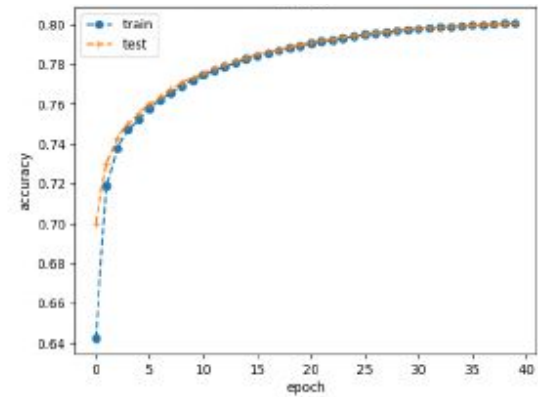
Identificación de fotones con CNN

Contraste de entrenamiento en función de las épocas de entrenamiento entre los dos modelos según la precisión (arriba) y según la función de pérdida binnary cross entropy (abajo).

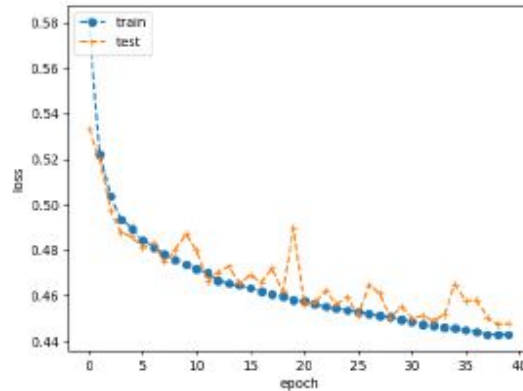
En todos los casos mostramos para cada gráfico los cálculos tanto para el entrenamiento (train puntos azules) y la validación (test cruces naranjas).



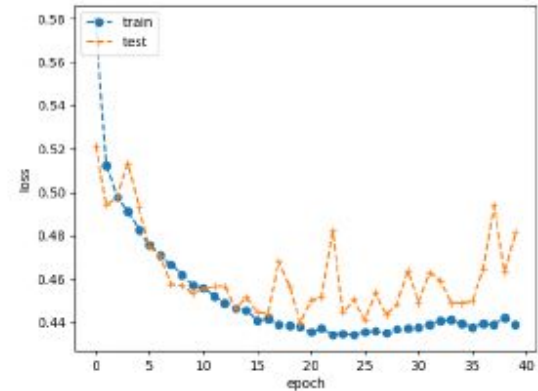
(a) modelo 0



(b) modelo 1

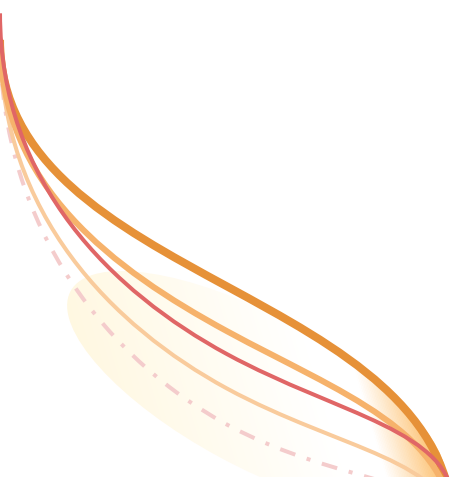


(c) modelo 0



(d) modelo 1

Resultados

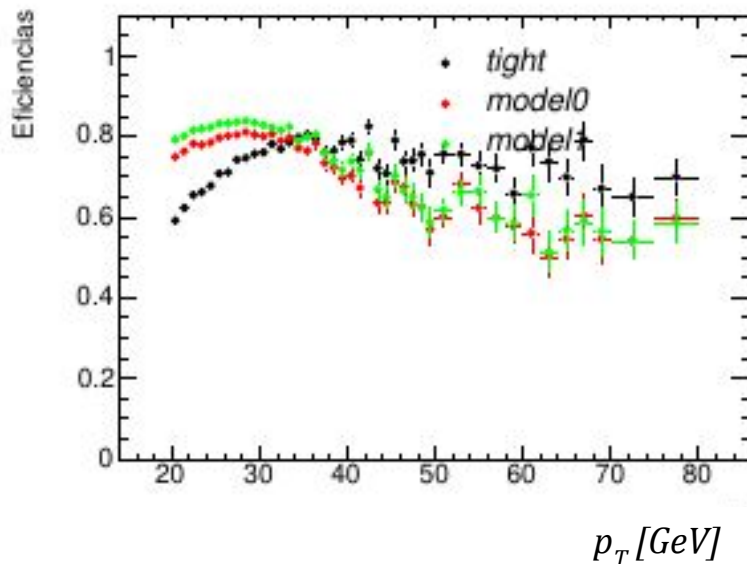
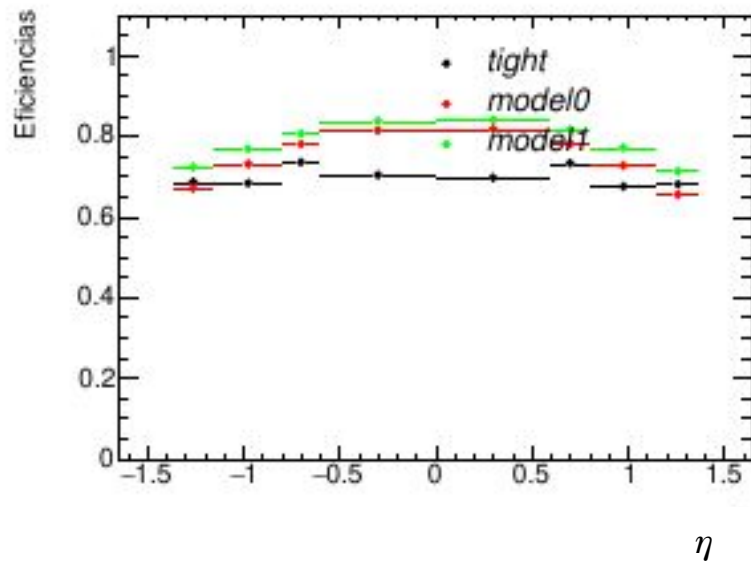


Resultados

$$\epsilon = \frac{N(x)^{pass}_{data}}{N(x)^{total}_{data}}$$

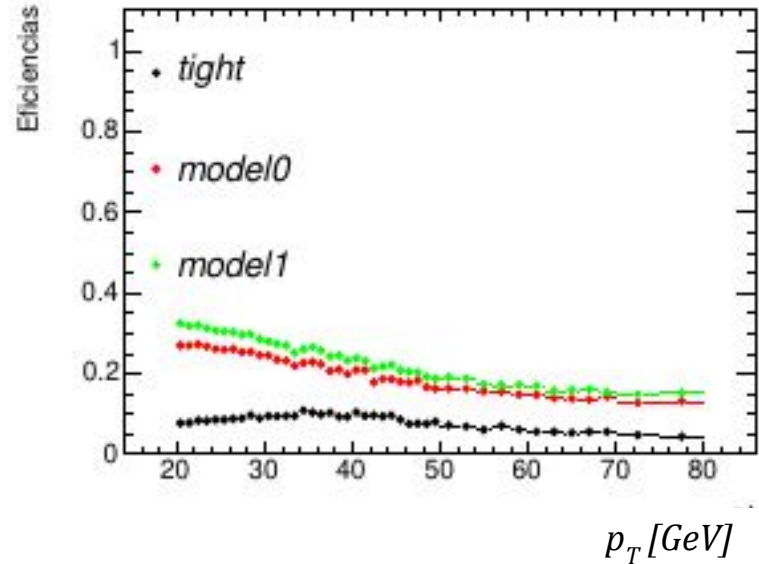
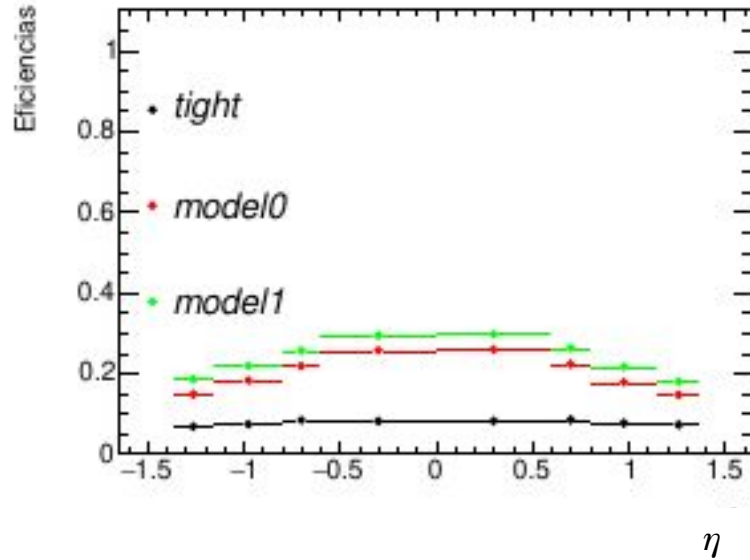
- $N(x)^{total}$: número total de eventos en la muestra completa.
- $N(x)^{pass}$: número de eventos en el subconjunto en el cual el fotón pasa los cortes del análisis.
- x es el observable medido.

Resultados



Comparación de eficiencias entre los métodos tight, modelo 0 y modelo 1 de CNN en eventos de señal. Las eficiencias se muestran en función de la pseudorapidez η (izquierda) y del momento transversal p_T en GeV (derecha).

Resultados



Comparación de eficiencias entre los métodos *tight*, *modelo 0* y *modelo 1* de CNN en eventos de fondo. Las eficiencias se muestran en función de la pseudorapidez η (izquierda) y del momento transversal p_T en GeV (derecha).

Conclusiones

Conclusiones

- Desarrollamos un paquete de software que construye una imagen a partir de la energía depositada en las celdas de la segunda capa del calorímetro electromagnético de los topo-clusters de los candidatos fotones en la zona central de pseudorapidez $|\eta|$ del detector (barril).
- A partir de los datos de ATLAS, utilizamos el proceso de decaimiento radiativo del bosón Z para obtener una muestra no sesgada de señal y un corte de inversión de aislamiento para obtener una muestra enriquecida de fondo.
- Utilizamos estas muestras para alimentar y entrenar un conjunto de dos modelos de redes neuronales convolucionales con el objetivo de que aprenda a clasificar fotones reales de falsos fotones (fondo) a partir de las imágenes construidas de la energía depositada en las celdas de los topo-clusters. Utilizando 60.000 eventos de ATLAS colectados en 2018, obtuvimos un 80 % de precisión en el entrenamiento de la red.

Conclusiones

- Con ambos modelos entrenados calculamos la eficiencia en la identificación de fotones y en el rechazo de fondo de éstos, y la contrastamos con las calculadas utilizando la identificación actual de cortes rectangulares.
- Observamos que tanto en función de η como de p_T , la **identificación *tight* provee un mejor rechazo de fondo**, marcando una diferencia entre los métodos que ronda valores superiores al 10 %. Mientras que en la identificación de señal, **ambos modelos de CNN presentan mayor eficiencia en la selección de fotones reales en η** .
- Como resultado de este trabajo hemos desarrollado un conjunto completo de herramientas de análisis que permite implementar una identificación de fotones sin utilizar cortes rectangulares, que dado que se realiza sólo a partir de la energía depositada en la segunda capa del calorímetro electromagnético es posible de utilizar en la etapa de reconstrucción rápida del HLT.

Agradecimientos

¡Muchas gracias!

