Development of Machine Learning PID Methods in KamLAND-Zen Experiment

Tohoku U. RCNS

Jun Nakane, KamLAND-Zen collaborations

NPML 2025 (10/28)

1. Background

- Neutrinoless double beta decay ($0\nu\beta\beta$)
- KamLAND-Zen experiment
- Particle identification (PID)

2. Datasets and models for PID

- Datasets
- KamNet / ViViT/ KamNet-ViViT / S2ViViT

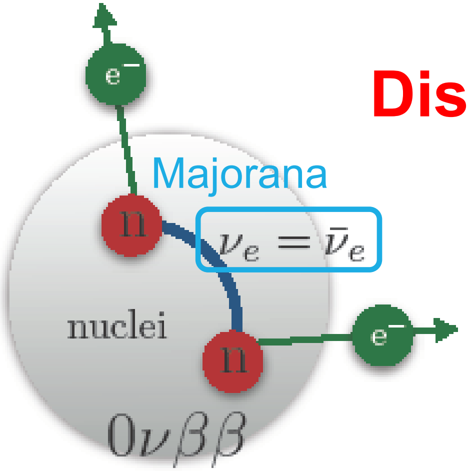
3. Development of integrated model

- KamNet vs. ViViT
- Performance of KamNet-ViViT / S2ViViT

4. Summary and prospects

Neutrinoless Double Beta Decay ($0\nu\beta\beta$)

Neutrinoless Double Beta Decay ($0\nu\beta\beta$)



Discovery of $\boldsymbol{0\nu\beta\beta}$

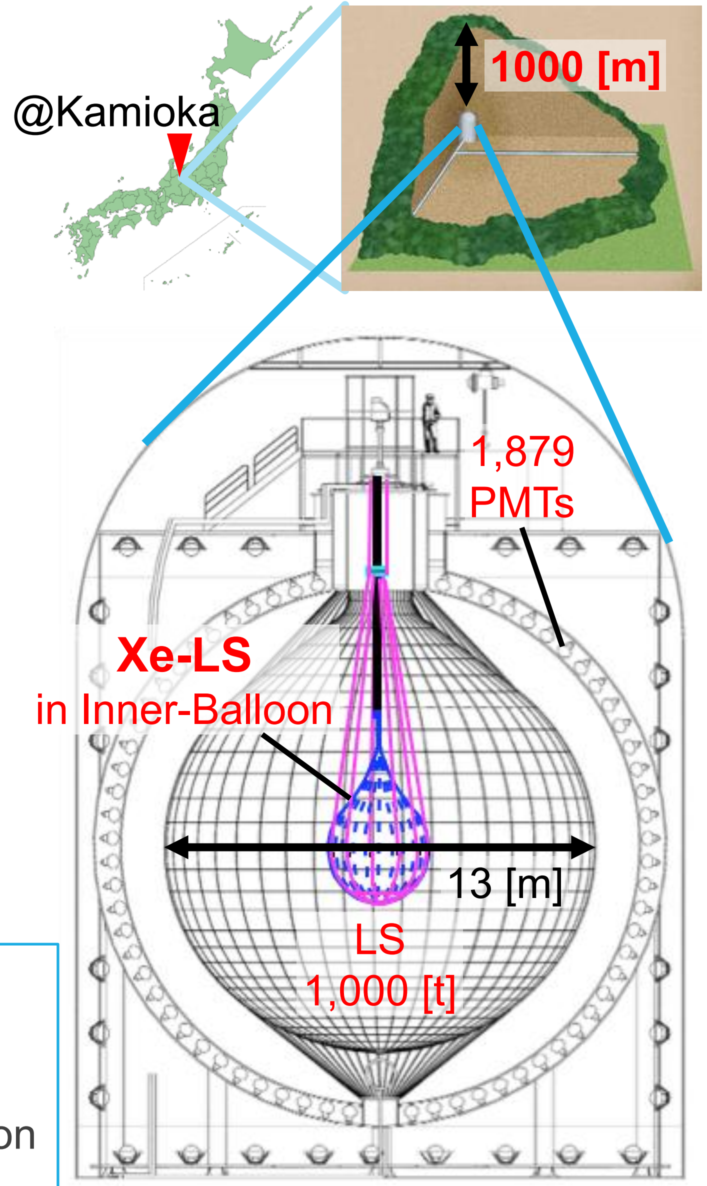
→ Verification of Majorana neutrino
Proof of beyond-Standard-Model physics
Elucidation of matter-dominated universe

KamLAND Detector

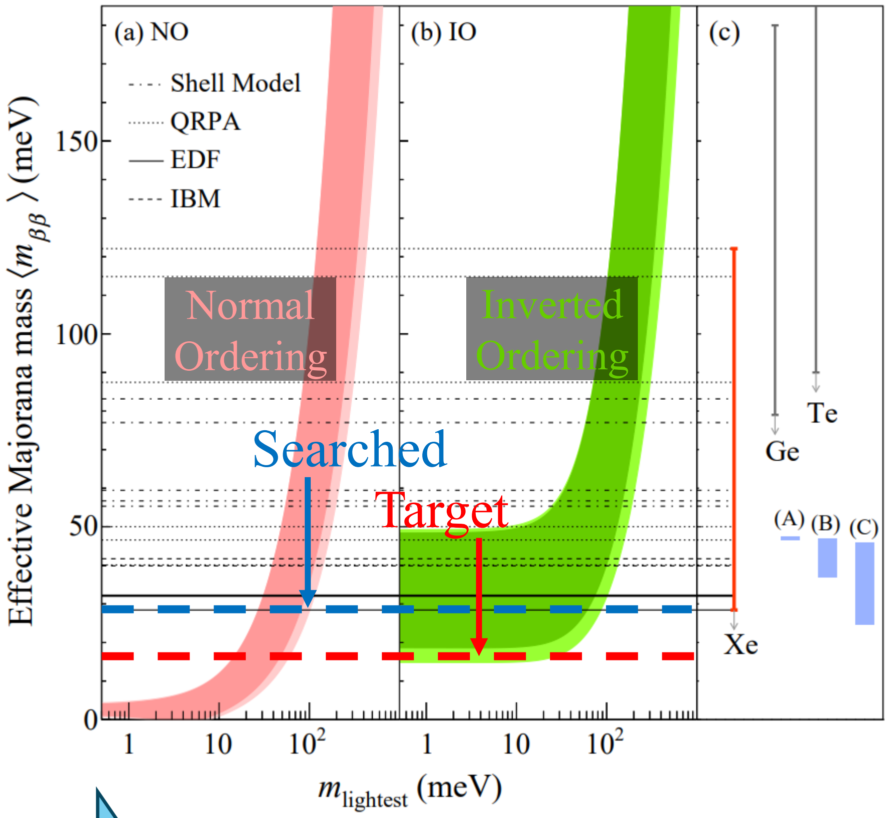
- **High-sensitivity** (energy resolution σ: 4% $(E = 2.5\ \mathrm{MeV})$)
- Extremely **low-radioactivity** $(\mathrm{U, Th} < 10^{-17}\ [\mathrm{g/g}])$

→ **Ideal detector for rare decay search!**

^{136}Xe: Double beta decay source
- $0\nu\beta\beta$ Q-Value: 2.46MeV (below ^{208}Tl γ BG)
- Long $2\nu\beta\beta$ half life
- (Relatively) easy to enrich/purify by distillation
- Dissolved into liquid scintillator (LS) at 3%



KamLAND-Zen Experiment



KamLAND-Zen experiment for $0\nu\beta\beta$ search

= ^{136}Xe + KamLAND

- Setting lower limit of half-life of $0\nu\beta\beta$ from the observation period and the amount of Xe

→ Setting upper limit of effective Majorana mass

$$\left(T_{1/2}^{0\nu}\right)^{-1} = G^{0\nu}|M^{0\nu}|^2\langle m_{\beta\beta}\rangle^2$$

($G^{0\nu}$: Phase space factor, $|M^{0\nu}|^2$: Nuclear matrix element (NME))

➔ Determination of neutrino mass ordering

Result

2011～2015

KamLAND-Zen 400

(Xe: $320{\sim}380$ [kg])

Increase in ^{136}Xe

2019～2024

KamLAND-Zen 800

(Xe: 745 ± 3 [kg])

Now Updating!!

2027~

KamLAND**2-Zen**

(Xe: $\sim 1{,}000$ [kg])

Result(400+800): First in the world to reach IO band

$$T_{1/2}^{0\nu} > 3.8 \times 10^{26}\ \text{[years]},$$

$$\langle m_{\beta\beta}\rangle < 28 - 122\ \text{meV}$$

arXiv 2406.11438 (2024)

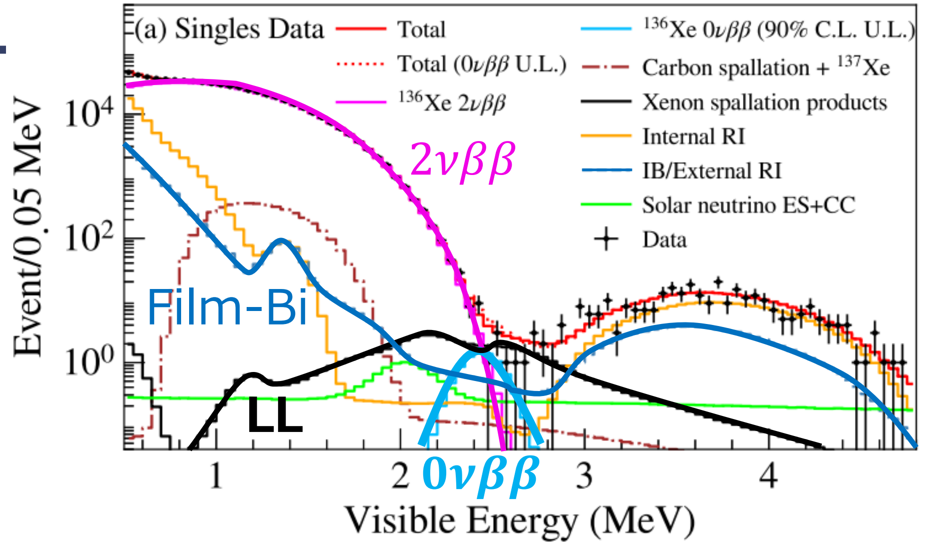
Target: Covering most of IO

Toward $T_{1/2}^{0\nu} = 2.0 \times 10^{27}$[yesrs],

$$\boldsymbol{\langle m_{\beta\beta}\rangle = 15\ \text{meV}}$$

Background in KamLAND-Zen

Major Background for $0\nu\beta\beta$



- $2\nu\beta\beta$, Solar ν

  → Distinguishable only by improved energy resolution

  ➔ **Detector upgrade** in progress (light yield x5)

- **long-lived ^{136}Xe spallation products via muon interaction** (LL)
  XeLS / **Inner Balloon** containing ^{238}U、^{232}Th derived ^{214}**Bi** (XeLS/Film-Bi)
  XeLS containing ^{12}C derived ^{10}C

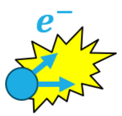
  → Reducible by improved energy/vertex resolution
  developed analysis methods

  ➔ Development of **Machine Learning PID Methods** (this talk)

Particle Identification (PID)

$0\nu\beta\beta$

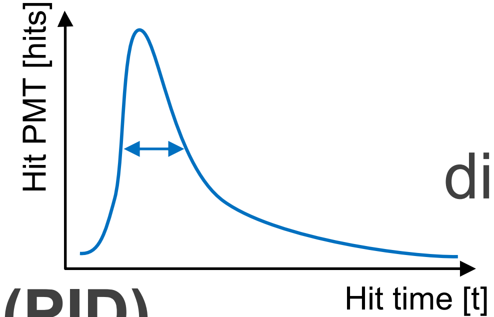
only β decay

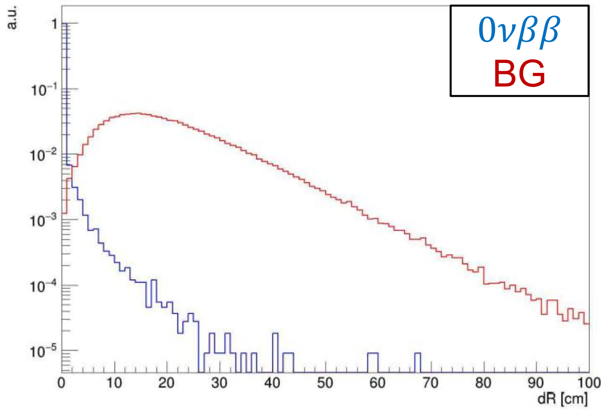


(~1cm)

‖

Narrow width





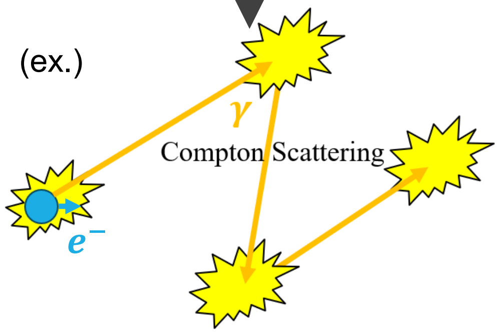
Emission distribution (MC sim)

Hit time distribution

BG(LL, Bi, ^{10}C)

β decay + γ **decay**

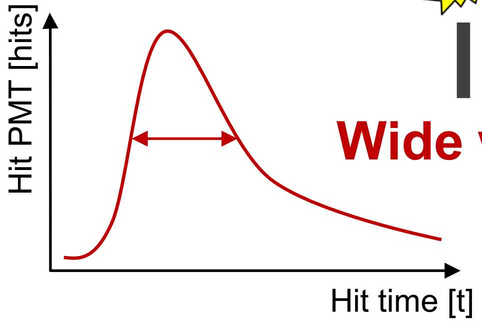
(ex.)



(~20cm)

‖

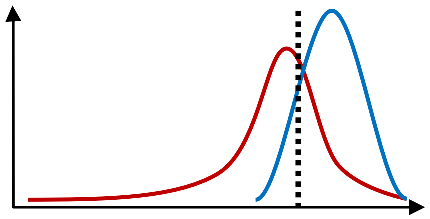
Wide width



Particle Identification (PID)

Input data PMT Hit Map → Neural network → Output data PID Score → Score distribution → **Signal** or **BG** **Identify**

(= Spatiotemporal (Hit time × PMT position) data
= Tensor: $28 \times (38 \times 38)$)



1. Background
    - Neutrinoless double beta decay ($0\nu\beta\beta$)
    - KamLAND-Zen experiment
    - Particle identification (PID)

2. Datasets and models for PID
    - Datasets
    - KamNet / ViViT/ KamNet-ViViT / S2ViViT

3. Development of integrated model
    - KamNet vs. ViViT
    - Performance of KamNet-ViViT / S2ViViT

4. Summary and prospects

Datasets

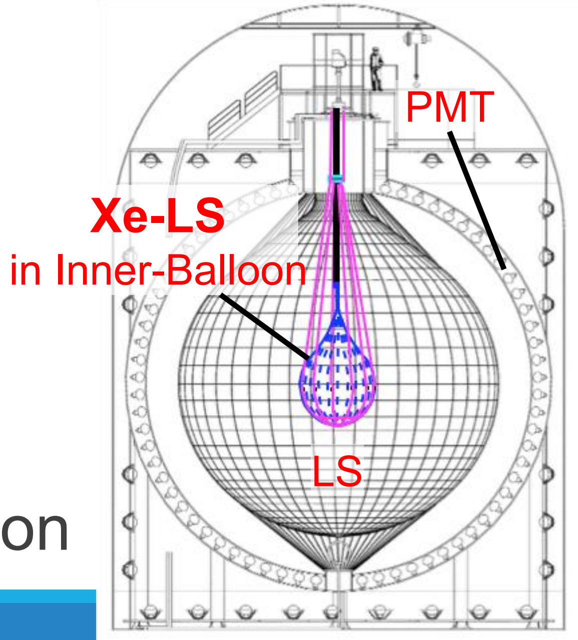
	Signal	BG
Train	XeLS-**Flat**Beta	XeLS-**Flat**Gamma (XeLS-**Flat**BetaGamma)
test	XeLS-$0\nu\beta\beta$	C10 LL-ALL Nuclear LL-ROI Nuclear **Bi214-XeLS**, Bi214-Film **Bi214-Data**

MC simulation data

Real data (Selected by coincidence for sequential decay)

Flat data: Uniform energy/vertex events within XeLS

Nuclear data: Each energy distribution/ uniform vertex events within XeLS/on the IB surface



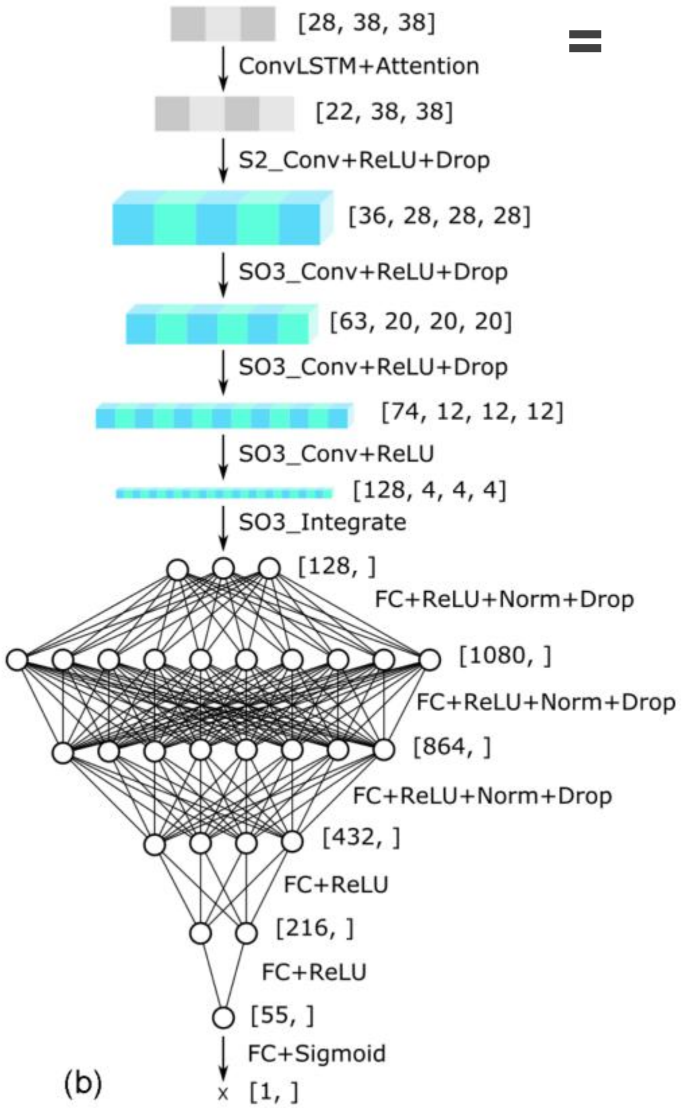
※ BG for Training: Although FlatBetaGamma is simulation based on actual events, FlatGamma yields improved rejection performance. Therefore, this talk shows results learned using FlatGamma.

Signal for test: Comparing Bi214-XeLS and Bi214-Data allows us to verify the practicality of the method.

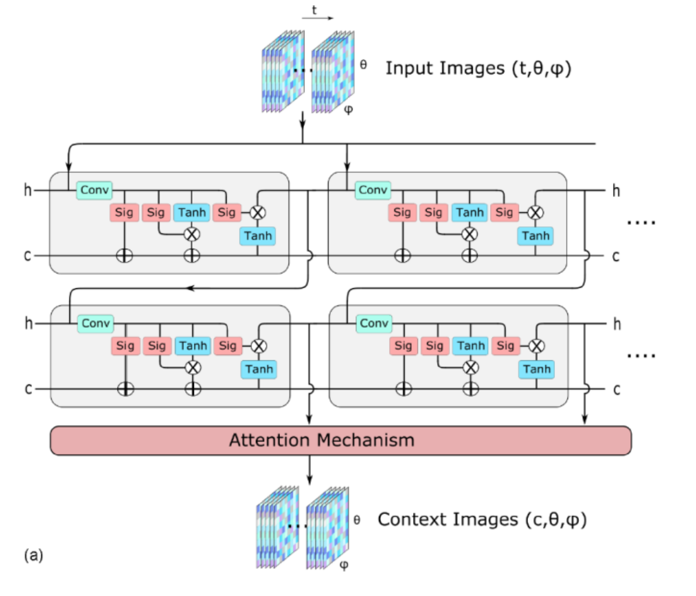
→ BG rejection efficiency: Rate of BG rejected by 10% signal rejection

KamNet (based on CNN) DOI: 10.1103/PhysRevC.107.014323

KamNet: CNN model developed specifically for experiments using KamLAND



= ConvLSTM

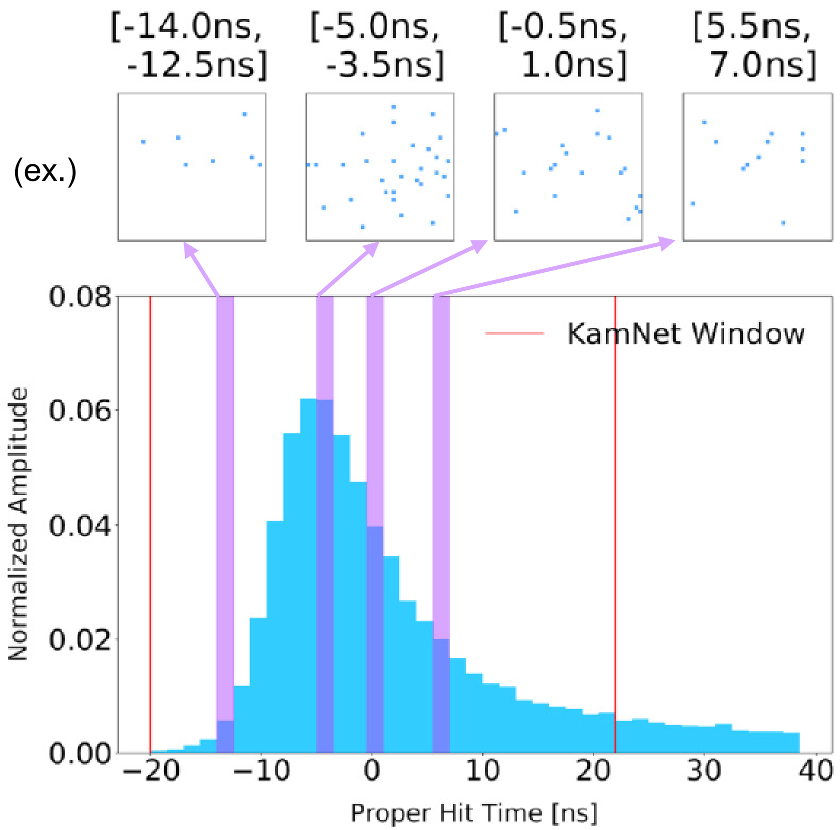


\+ Spherical symmetric CNN (S2CNN)

Consider spherical symmetry of KamLAND

- Input data: PMT Hit Map = Spatiotemporal (Hit time × PMT position)
- CNN: Focusing on **local** features Resistant to **low**-frequency noise

(ex.)



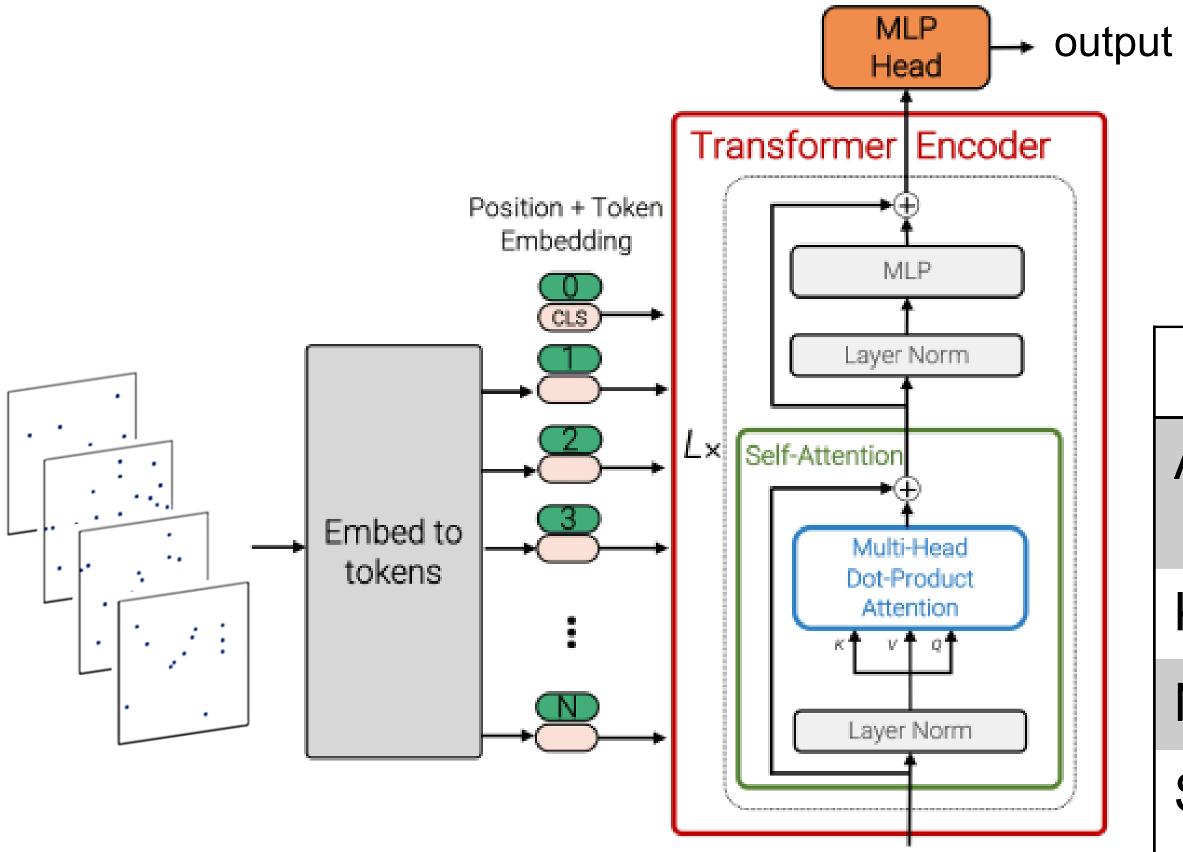
Video Vision Transformer: ViViT

DOI: 10.48550/arXiv.2103.15691

Vision Transformer:

Image recognition model using attention mechanism

➔ ViViT: Model extended to video



- Features of KamNet

  Limited to local features due to CNN

  Vulnerability to high-frequency noise

- Benefits of ViViT

  Global feature extraction

  Focus on critical spatiotemporal domains by AM

	KamNet	ViViT
Architecture	ConvLSTM + S2CNN	Video Vision Transformer
Key Features	Local pattern	Global pattern
Noise Resistance	Low-frequency	High-frequency
Symmetry	Explicitly consider spherical symmetry	Implicit learning in AM

Development of Combined Model

- KamNet-ViViT = KamNet score + ViViT score

  Scores generated by individual training and testing are added up event by event, and rejection efficiency is calculated based on total score

Input → KamNet process → KamNet score
Input → ViViT process → ViViT score
KamNet score + ViViT score → Output PID score

- S2ViViT = ViViT + S2CNN

  Model enhanced for spherical symmetry by combining S2CNN layer after ViViT processing

Input → ViViT process → S2CNN layers → fully-connected layer → Output PID score

1. Background
- Neutrinoless double beta decay ($0\nu\beta\beta$)
- KamLAND-Zen experiment
- Particle identification (PID)

2. Datasets and models for PID
- Datasets
- KamNet / ViViT/ KamNet-ViViT / S2ViViT

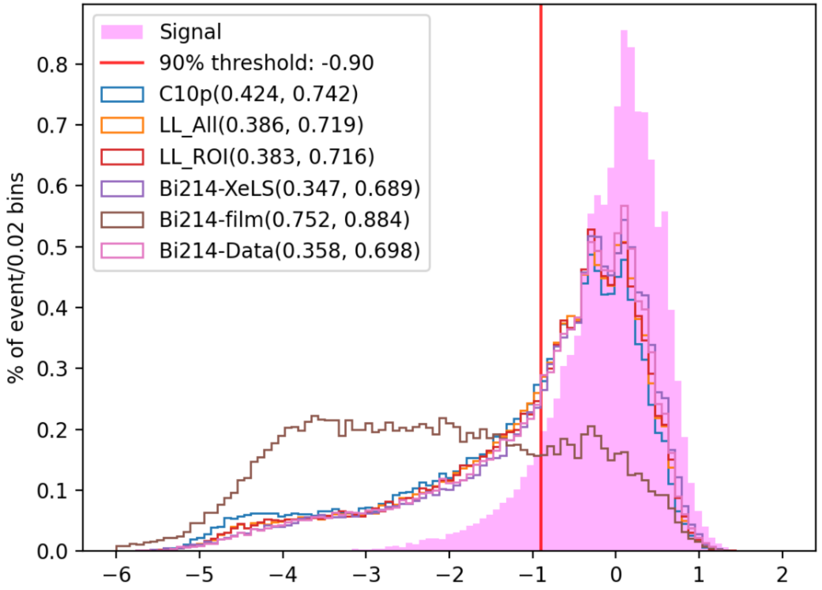
3. Development of integrated model
- KamNet vs. ViViT
- Performance of KamNet-ViViT / S2ViViT

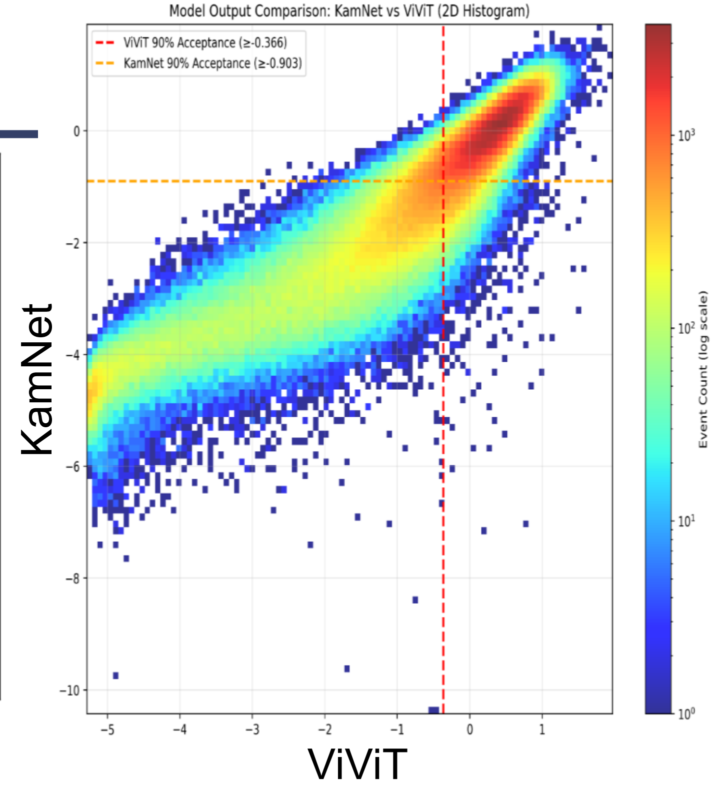
4. Summary and prospects

KamNet vs. ViViT

	KamNet	ViViT
$0\nu\beta\beta$	10%	10%
C10	42.4%	42.2%
LL-ALL	38.6%	38.3%
LL-ROI	38.3%	38.8%
Bi214-XeLS	34.7%	35.0%
Bi214-film	75.2%	70.5%
Bi214-Data	35.8%	36.9%

KamNet vs. ViViT: No significant difference



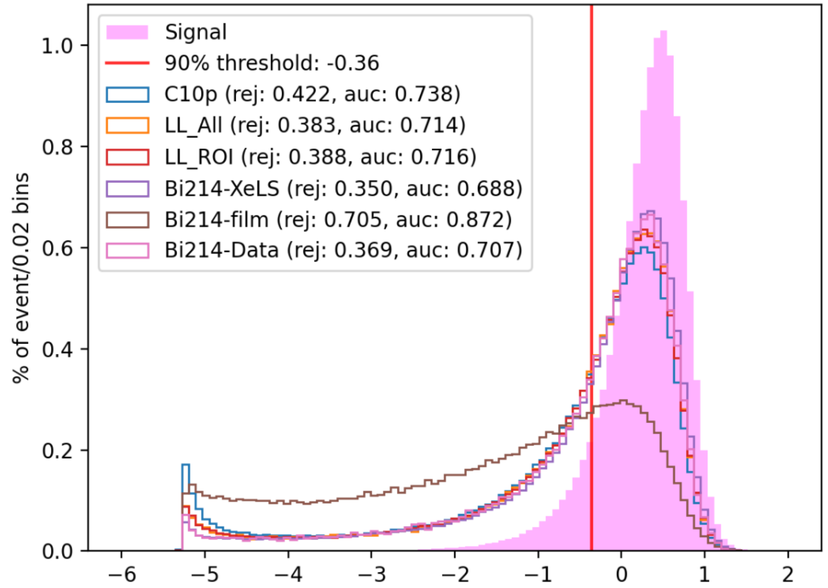


Correlation coefficient between both scores: ~0.9

→ Extracting not completely identical information (using different feature extraction mechanisms)

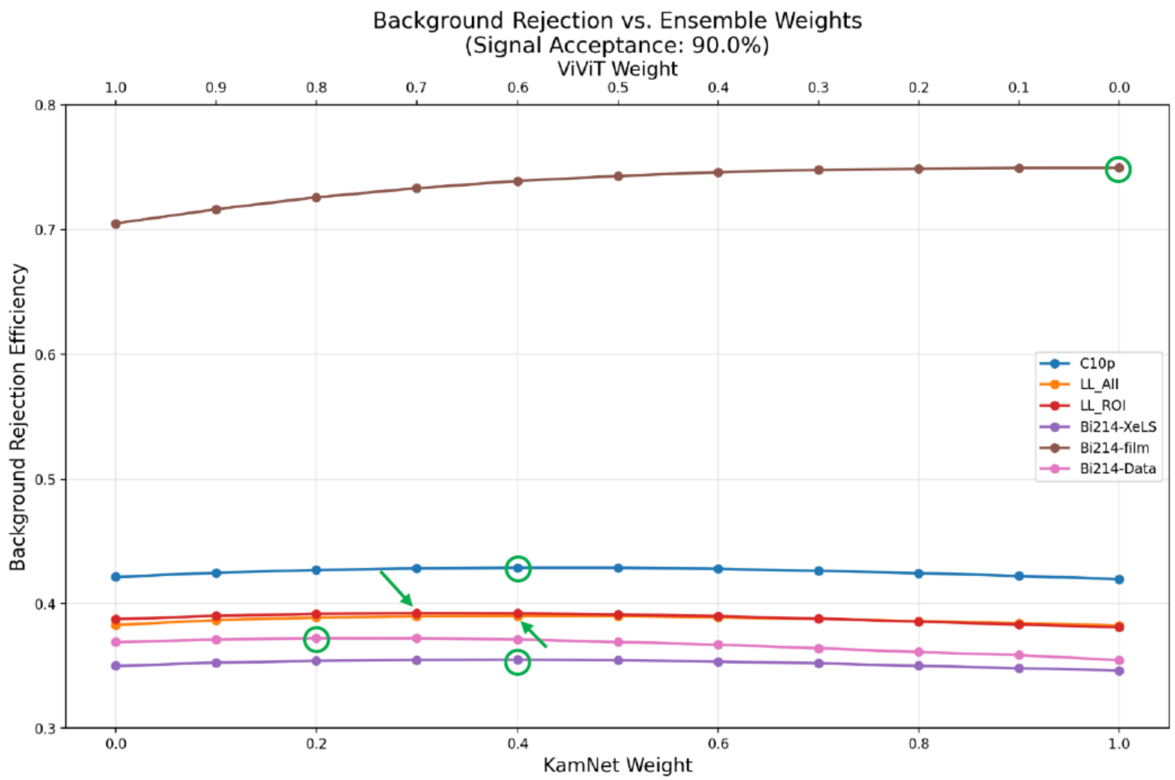
→ **Investigation of combined model**

(Performance improvement expected)



Performance of KamNet-ViViT

	KamNet	**KamNet-ViViT (4:6)**	ViViT
$0\nu\beta\beta$	10%	10%	10%
C10	42.4%	42.9%	42.2%
LL-ALL	38.6%	39.0%	38.3%
LL-ROI	38.3%	39.2%	38.8%
Bi214-XeLS	34.7%	35.5%	35.0%
Bi214-film	75.2%	73.9%	70.5%
Bi214-Data	35.8%	37.1%	36.9%



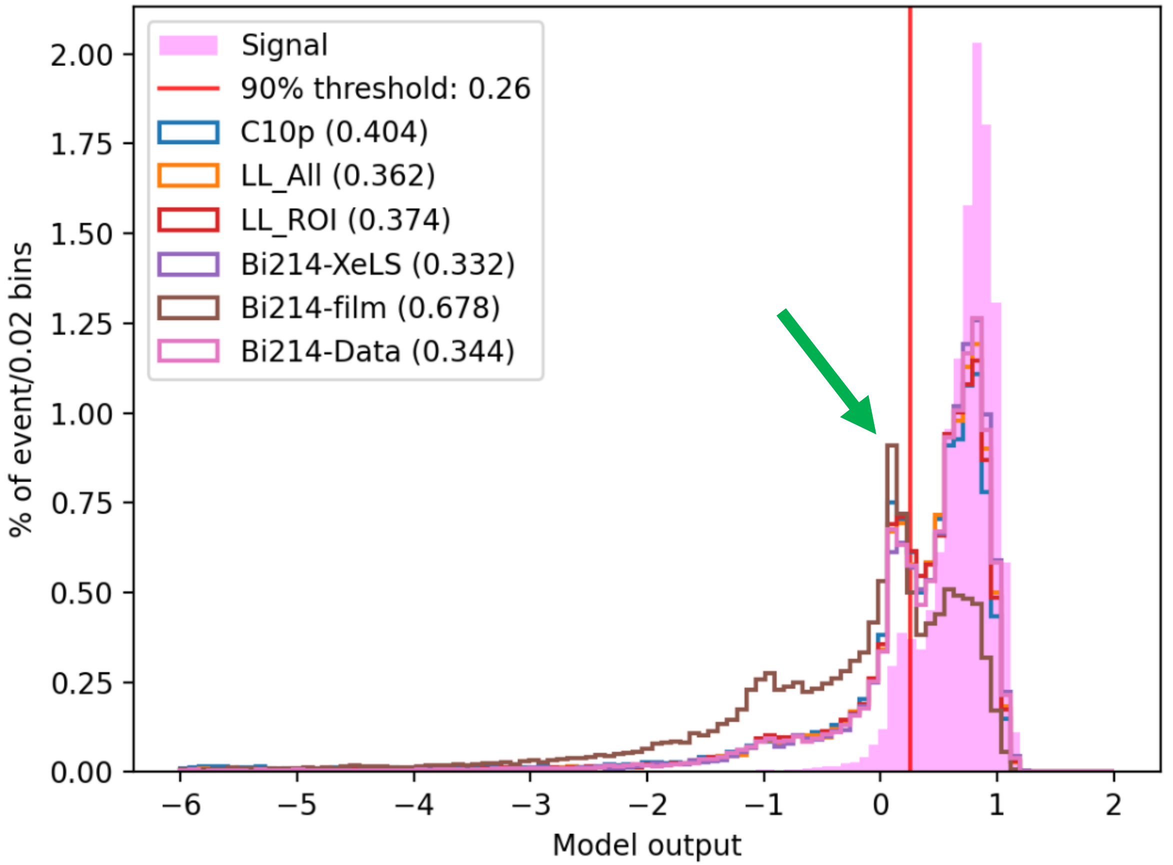
* Optimal weight for maximum efficiency in each BG

Compared to KamNet/ViViT,
it achieved improvement in performance (≃0.5%) (error: ±0.3%)

➔ **Contributes to improved sensitivity**
Increase in computer resources is within acceptable limits

Performance of S2ViViT

	KamNet	**S2ViViT**	ViViT
$0\nu\beta\beta$	10%	10%	10%
C10	42.4%	40.4%	42.2%
LL-ALL	38.6%	36.2%	38.3%
LL-ROI	38.3%	37.4%	38.8%
Bi214-XeLS	34.7%	33.2%	35.0%
Bi214-film	75.2%	67.8%	70.5%
Bi214-Data	35.8%	34.4%	36.9%



Compared to KamNet/ViViT,
it showed significant decrease in performance (≃2%) (error: ±0.3%)

Distinct peak appeared in BG
→ Optimizing signal rejection rate should improve rejection efficiency (maximize FoM)

1. Background
- Neutrinoless double beta decay ($0\nu\beta\beta$)
- KamLAND-Zen experiment
- Particle identification (PID)

2. Datasets and models for PID
- Datasets
- KamNet / ViViT/ KamNet-ViViT / S2ViViT

3. Development of integrated model
- KamNet vs. ViViT
- Performance of KamNet-ViViT / S2ViViT

4. Summary and prospects

Summary and prospects

- Objective of KamLAND-Zen: **Observation of $0\nu\beta\beta$** → Proof of Majorana Neutrino
- Hardware Improvements for Reducing BG: Updating KamLAND2 detector
  Software Improvements : **Developing PID using machine learning**
- Evaluating CNN-based KamNet and Transformer-based ViViT
  → Rejection efficiency is **comparable in performance**
  Extracting not completely identical information (using different feature extraction)
- Development and Evaluation of Model **Combining KamNet and ViViT**
  → ・**KamNet-ViViT** shows slight improvement in performance (≃0.5%)
  ・**S2ViViT** showed performance decline, but distinct peak appeared in BG

Prospects: Confirmation of uncertainty and statistical significance, Error assessment
Performance improvement by optimizing signal rejection rate
(maximizing FoM)
Consideration of **new input data/models**
(e.g., neutron distribution generated during nuclear fission/GNN)

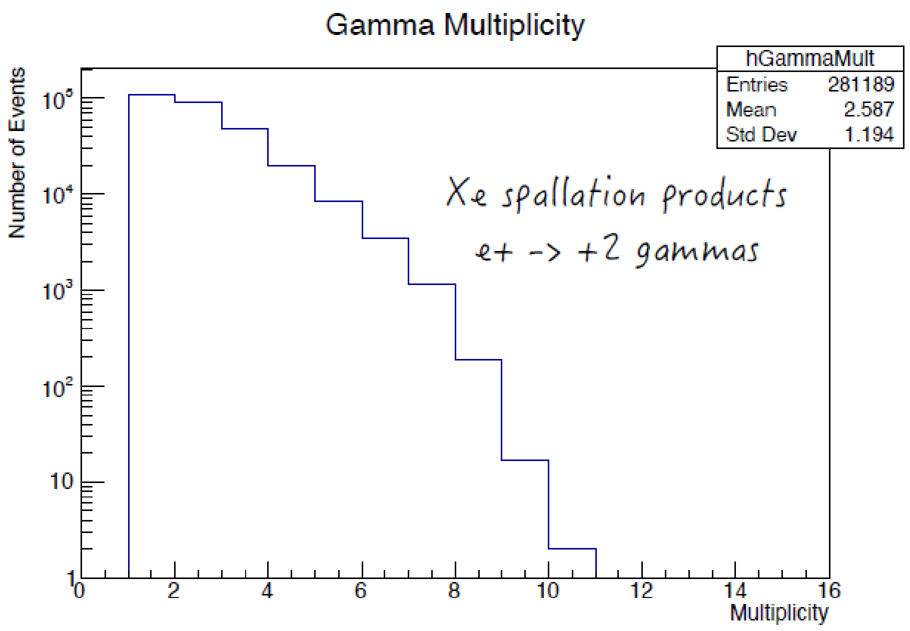
Back Up

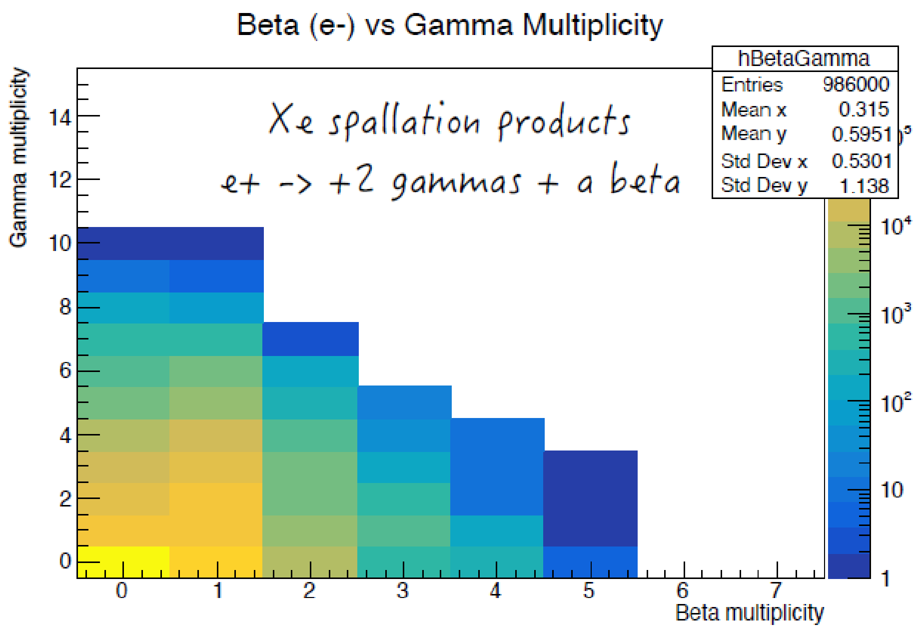
Data Set

Training datasets

For PID tasks, please use **FlatBetaForAIBM** as the signal sample and **FlatGammasForAIBM** as the background.
A radius cut (e.g., R < 160 cm) should be applied.

- **FlatBetaForAIBM**: Single beta decay events with flat energy distribution (0.5-8.5 MeV).
- **FlatGammasForAIBM**: Multi-gamma decay based on gamma multiplicity of Xe spallation products (~1000 nuclei), total energy uniformly distributed (0.5-8.5 MeV), each gamma randomly distributed.
- **FlatBetaGammaForAIBM**: Mixed multi-beta and gamma decay based on multiplicity of Xe spallation products.





4

Data Set

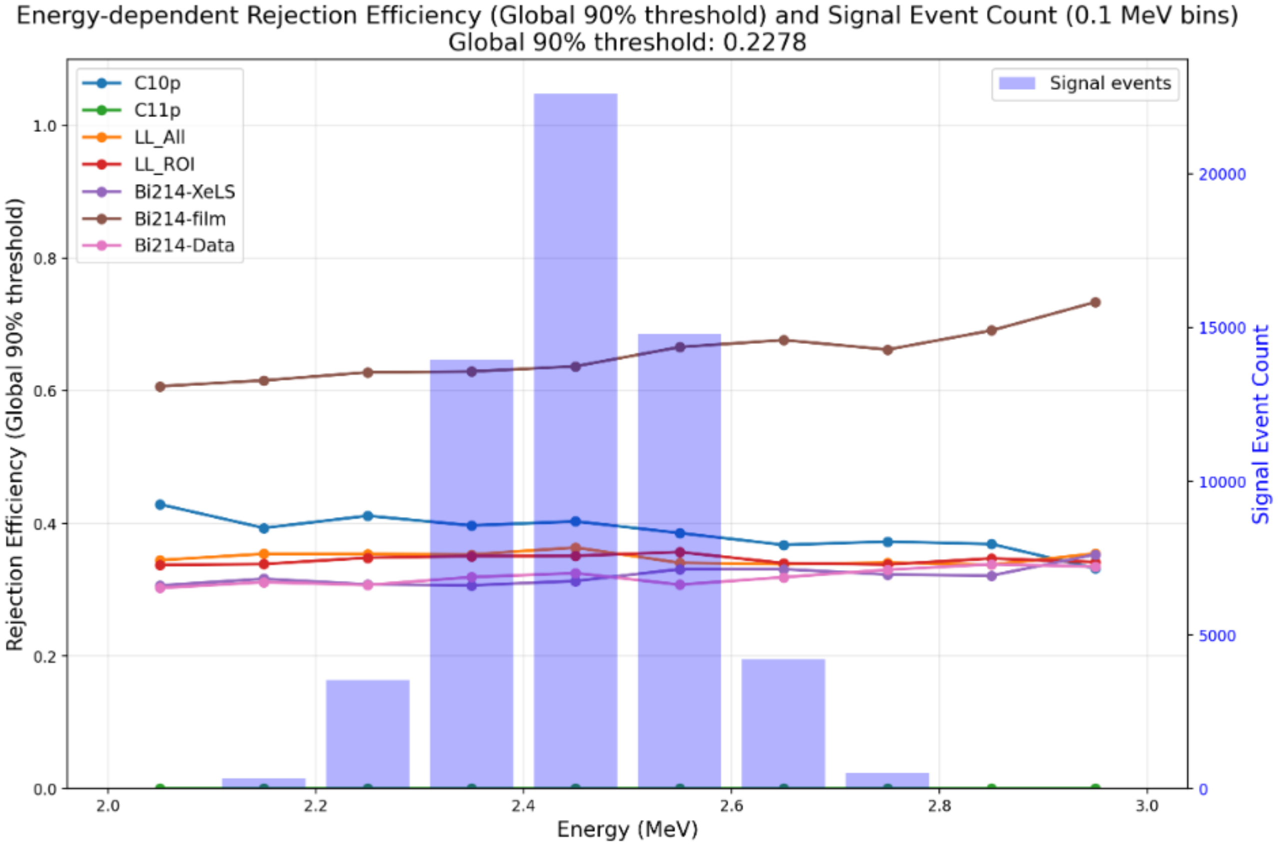
Validation datasets

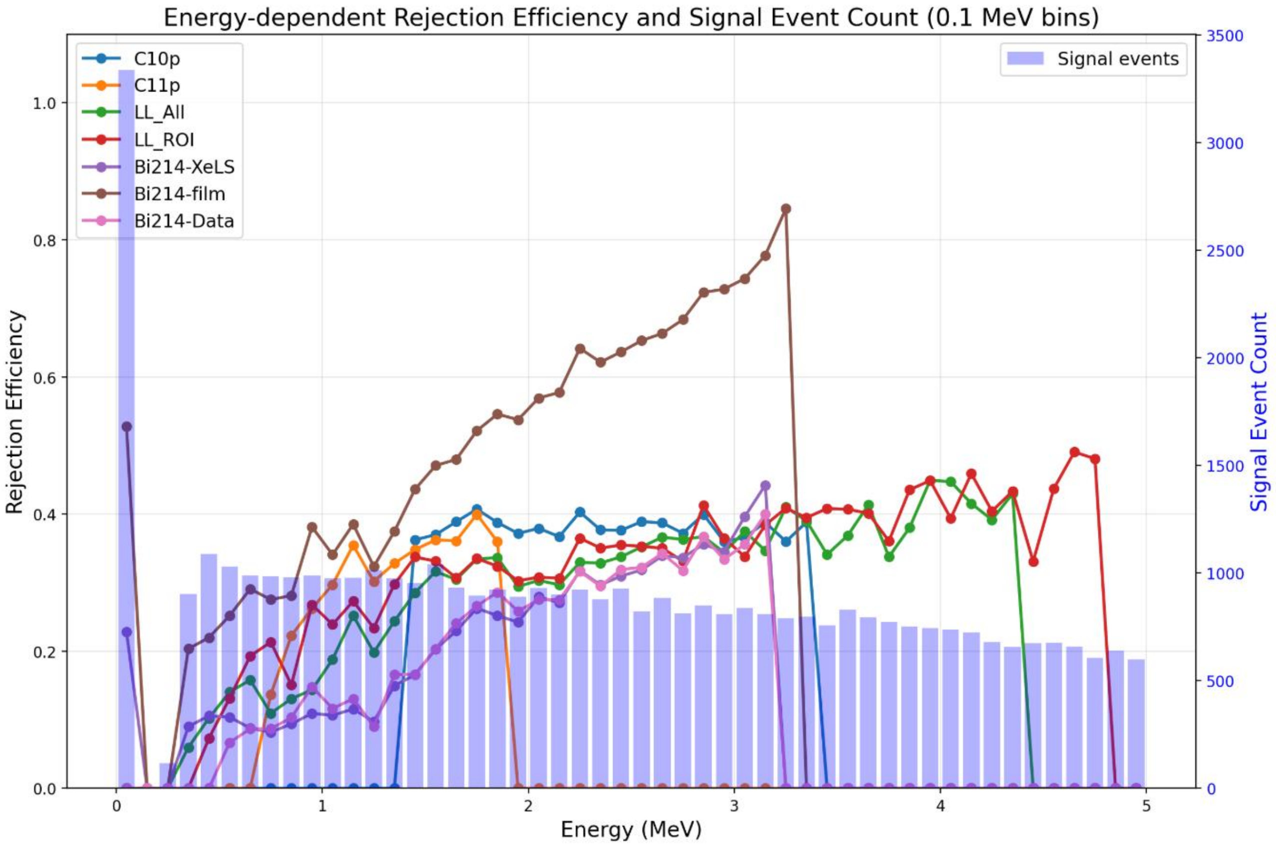
- **RealData-Bi214**: Real data with 250 cm fiducial volume cut.
- **XeLS-Bi214m**: Optimal for PID performance evaluation and data-MC agreement checks.
- **XeLS-C10p**: Optimal for O-Ps validation; real data may be provided soon (coming soon)
- **XeLS-C11p**: Optimal for low-energy PID performance evaluation
- **XeLS-LLmixROIForAIBM**: Major 32 Xe spallation isotopes with energy deposition in ROI. An o-Ps (orthopositronium) model is implemented based on recent measurements in XeLS (Funahashi, Master's thesis). For simplicity, a positron is modeled as **an electron with energy deposited at T=0**, and a **positron without energy at T=T_life.**
- **XeLS-LLmixAllForAIBM**: Xe spallation isotopes with energy deposition ≥0.5 MeV, mix ratio based on FLUKA and Geant4 RadioactiveDecay's rate predictions.

5

Validity of rejection efficiency

Signal: 0nu Solar





In the case of 0nu, the full range of BG rejection efficiency is calculated with 90% signal acceptance within a limited range. → Not calculating the correct efficiency

➔ **When developing the model, use solar as a flat beta signal**

ViViT (rejection efficiency depending on training dataset)

	+1 Gamma	+1 BetaGamma	+Charge Gamma	+Charge BetaGamma
0nu	90%	90%	90%	90%
C10	39.2%	35.1%	36.1%	33.0%
C11	47.1%	34.7%	40.6%	25.2%
LL-ALL	62.0%	55.1%	55.4%	41.6%
LL-ROI	39.8%	35.3%	36.3%	31.3%
Bi214-XeLS	32.6%	28.7%	30.2%	25.3%
Bi214-film	63.8%	56.5%	58.4%	50.0%
Bi214-Data	31.6%	27.2%	24.0%	21.4%

Best rejection efficiency

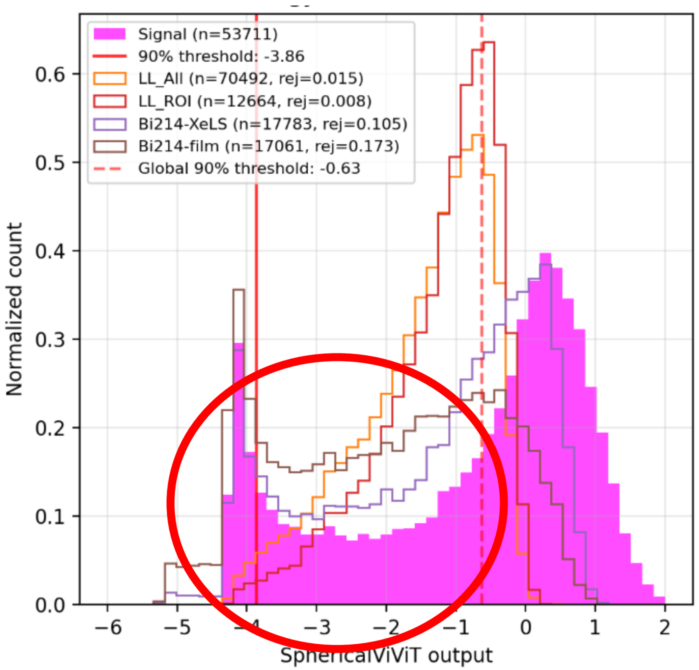
+1Gamma gives the best rejection efficiency (→ Reason is currently under investigation)

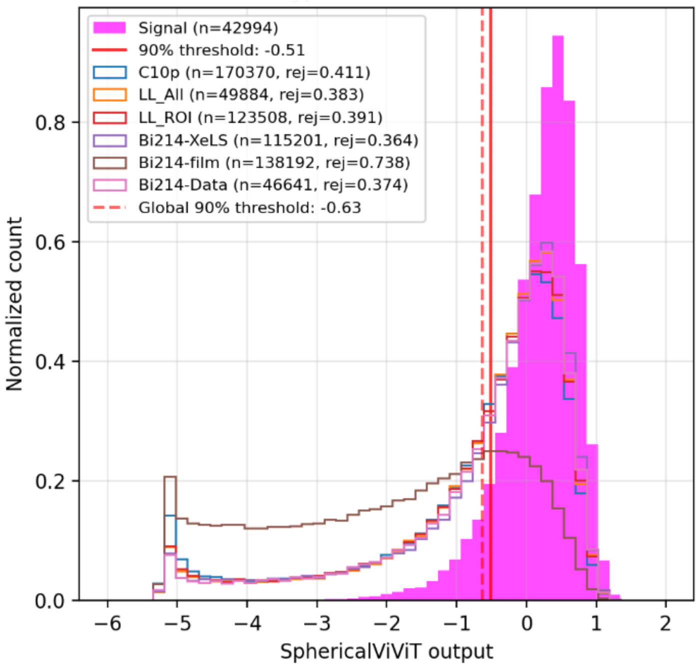
→ **When developing the model, use +1Gamma**

データセレクト

- モデルごとに十分学習するために必要なデータ数が異なる
  → 学習時間がおおよそ等しくなる学習用データ数を使用 (~15 minutes/epoch)
- 検証用信号の$0\nu\beta\beta$はROI(2.35~2.7 MeV)のみにイベントがある。
  → 検証用BGをROIのみにカット
- 低エネルギー領域(0.0~0.5 MeV)の識別能力が非常に悪い
  = 低エネルギーのデータが学習に悪影響を与えている
  → 学習用信号/BGの0.0~0.5 MeVのイベントをカット

Energy: **0.0-0.5 MeV** 2.35-2.7 MeV





Increasing the Amount of Training Data

	ViViT	ViViT	KamNet
data set	+1Gamma	+1Gamma	+1Gamma
[events]	150k →	1M	150k
Solar (Flat β)	90%	90%	90%
C10	32.7%	37.8%	36.5%
C11	25.2%	32.6%	28.5%
LL-ALL	42.4%	32.5%	35.4%
LL-ROI	31.3%	33.9%	33.2%
Bi214-XeLS	23.9%	27.1%	26.3%
Bi214-film	54.4%	61.5%	62.4%
Bi214-Data	23.4%	27.7%	26.6%

Transformer (ViViT) requires more data to train than CNN (KamNet)

✗ Aligning the amount of data
✓ Aligning the calculation times (~15 minutes/epoch)

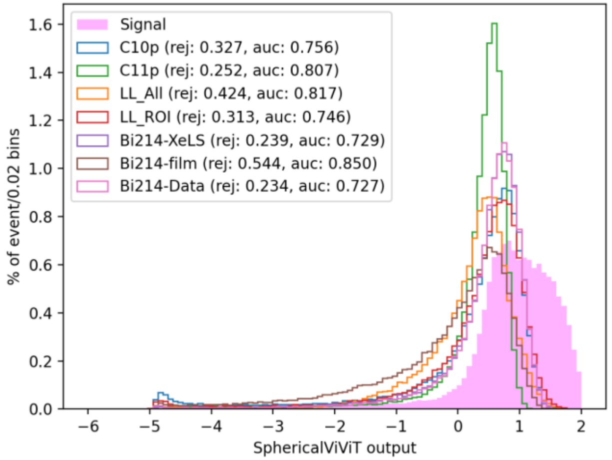
Rejection efficiency: the same or improved (except for LL-ALL)

Both KamNet and ViViT may be able to improve their efficiency by increasing the amount of data.

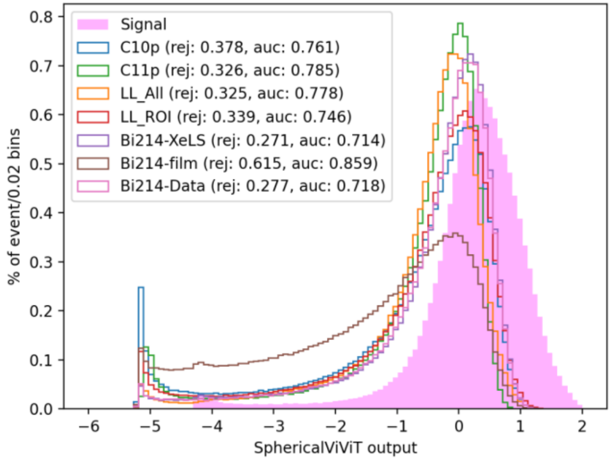
→ Verification is required.

Data Select

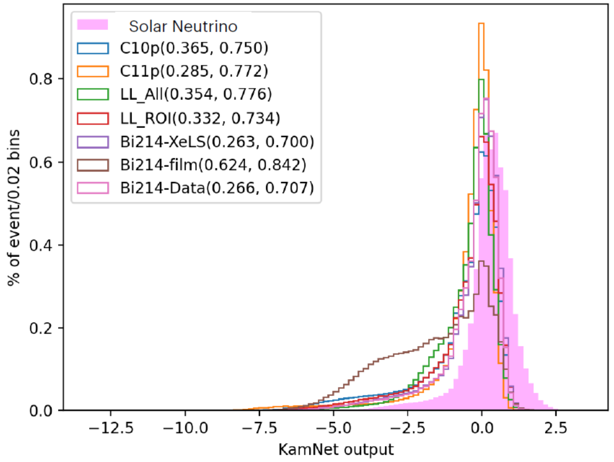
ViViT 150k



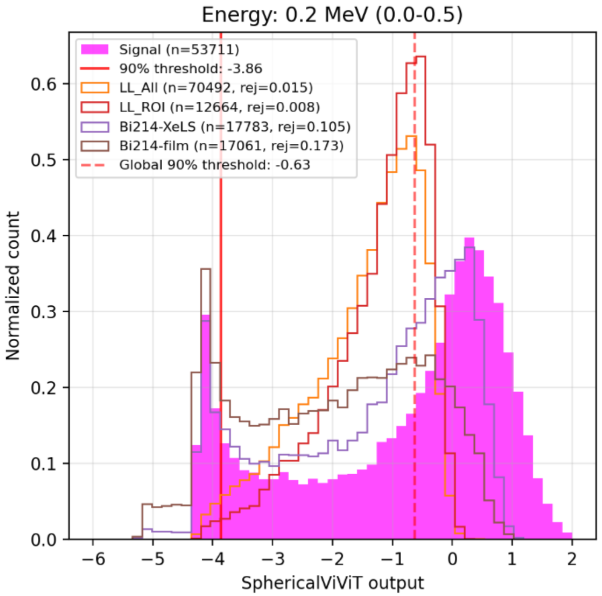
ViViT 1M



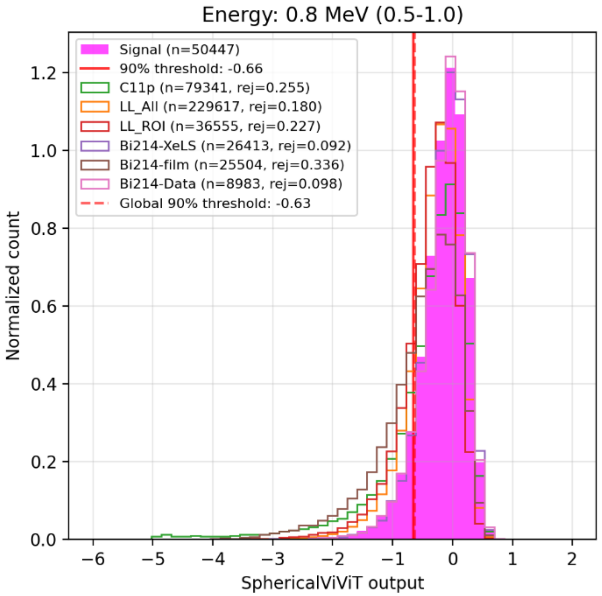
KamNet 150k



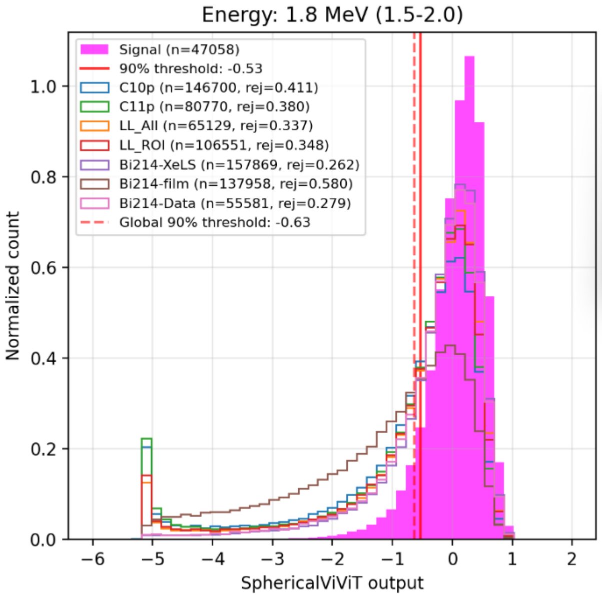
Energy: **0.0-0.5 MeV**



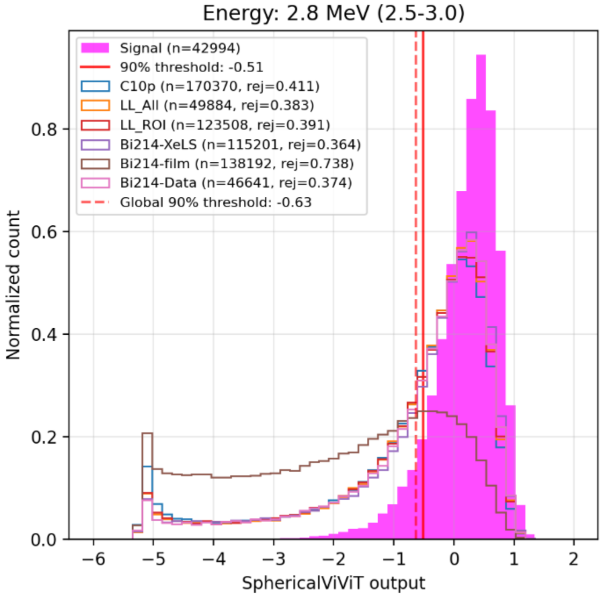
0.5-1.0 MeV



1.5-2.0 MeV



2.5-3.0 MeV

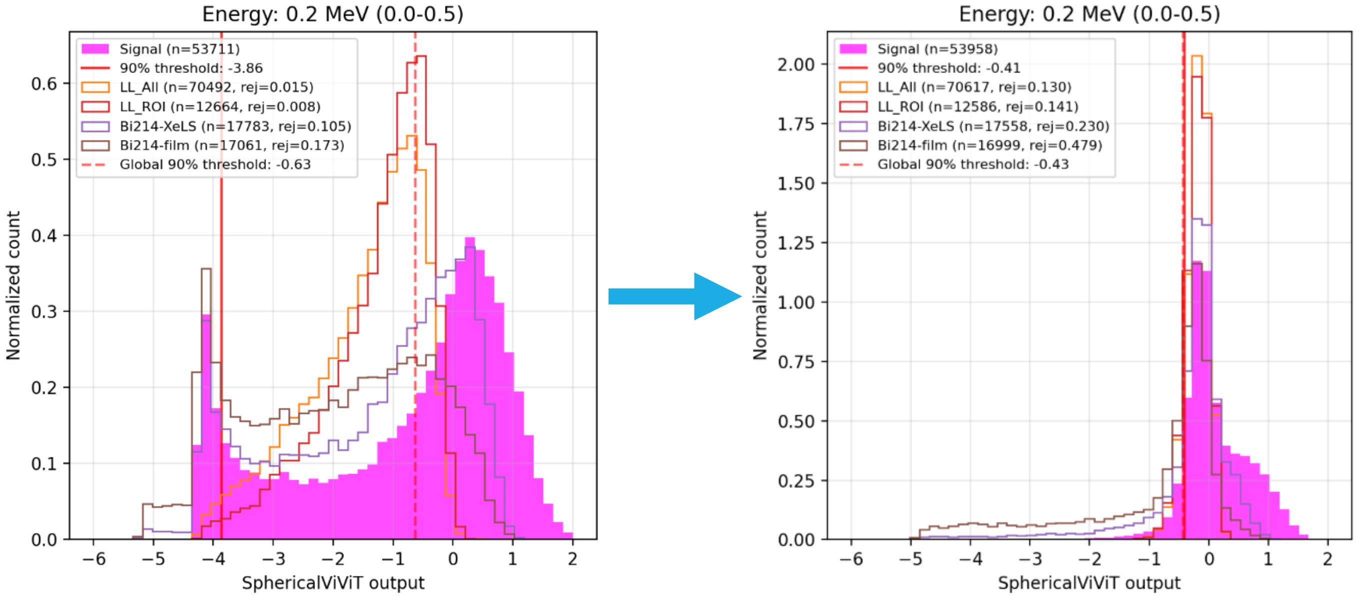


Energy Cut: > 0.5 MeV (1M events)

	ViViT	ViViT
data set	+1Gamma	+1Gamma
[events]	1M	1M
Energy Cut	Full range	> 0.5 MeV
Solar (Flat β)	90%	90%
C10	37.8%	48.2%
C11	32.6%	39.5%
LL-ALL	32.5%	26.8%
LL-ROI	33.9%	40.6%
Bi214-XeLS	27.1%	35.1%
Bi214-film	61.5%	68.5%
Bi214-Data	27.7%	36.5%

Output distribution (0.0-0.5 MeV)

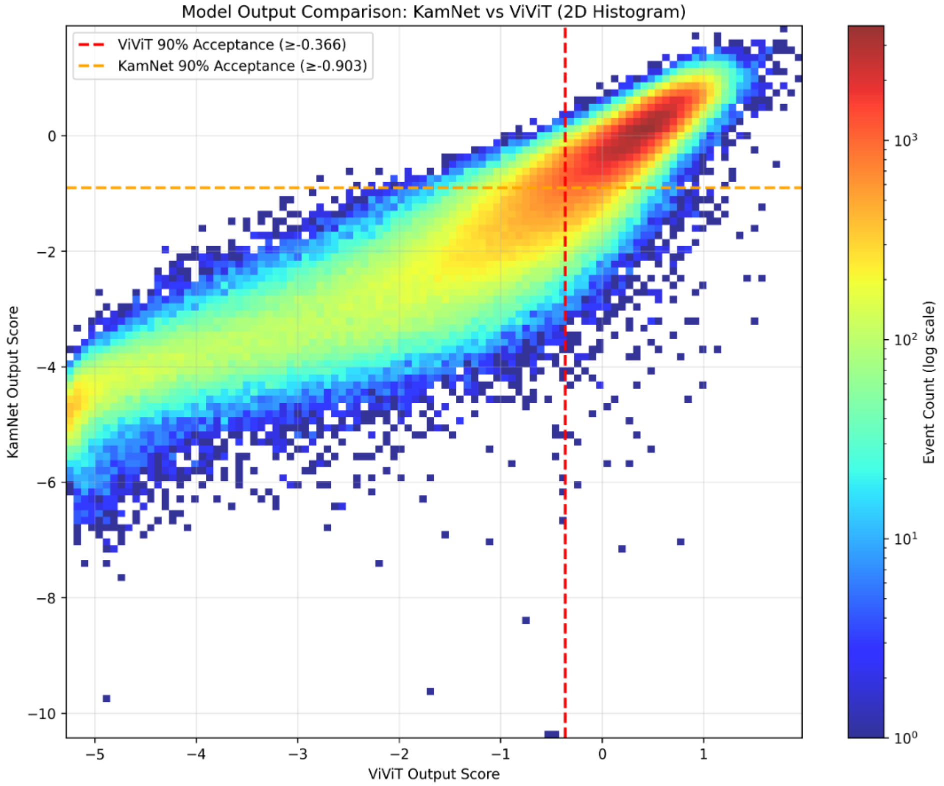
Full range > 0.5 MeV

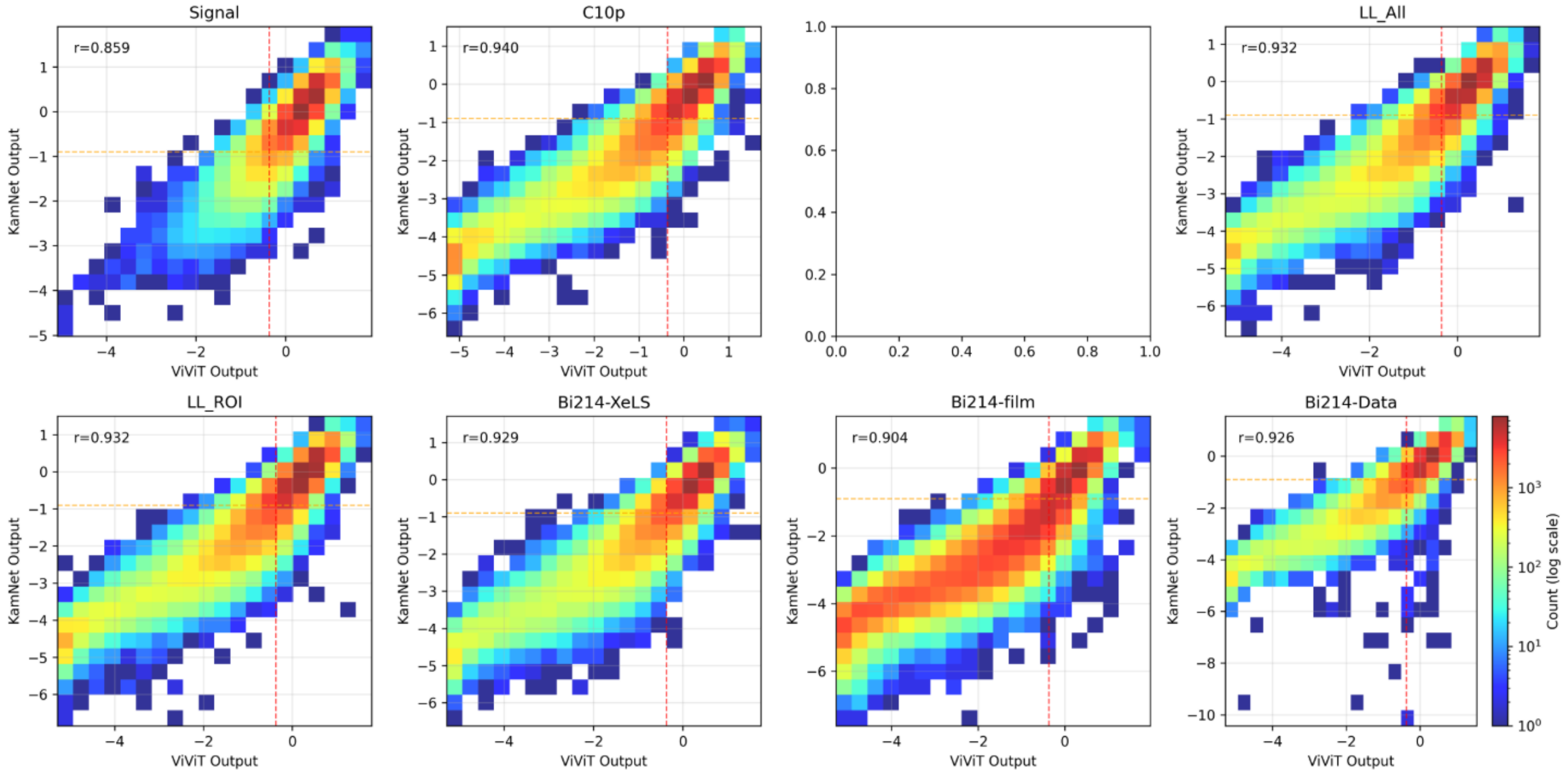


Improved identification ability for low energy (0.0-0.5 MeV)

Rejection efficiency is the improved, except for LL-ALL

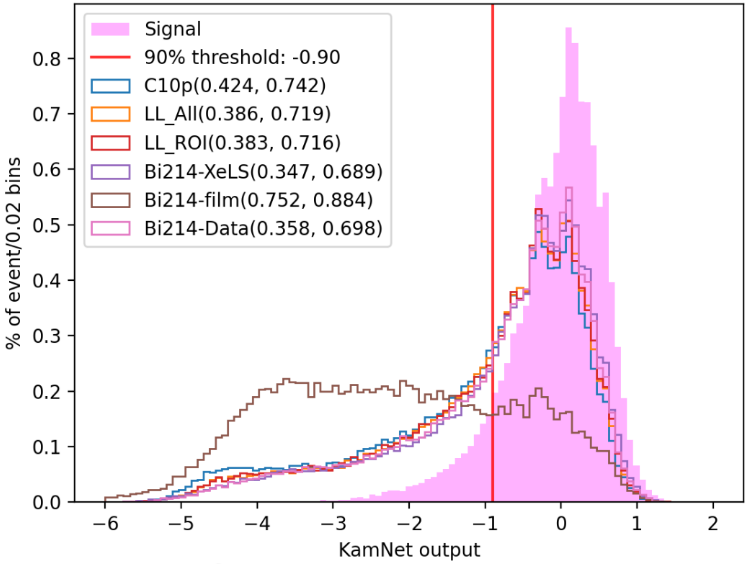
KamNet vs ViViT



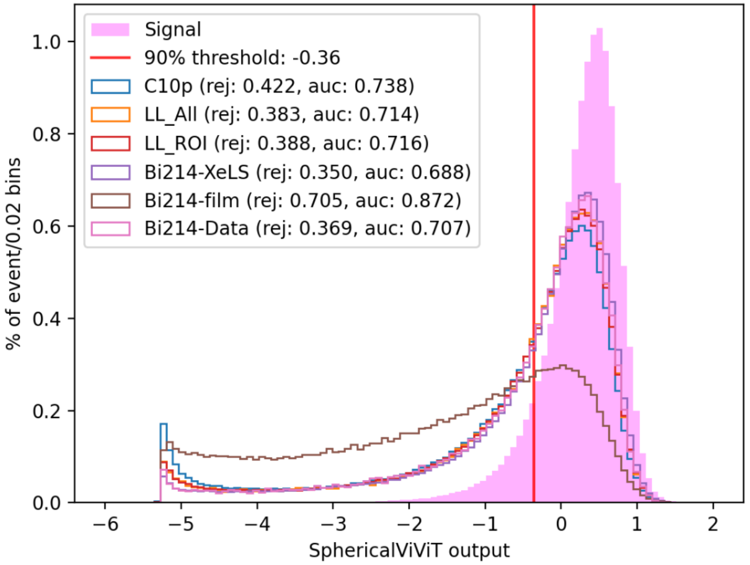


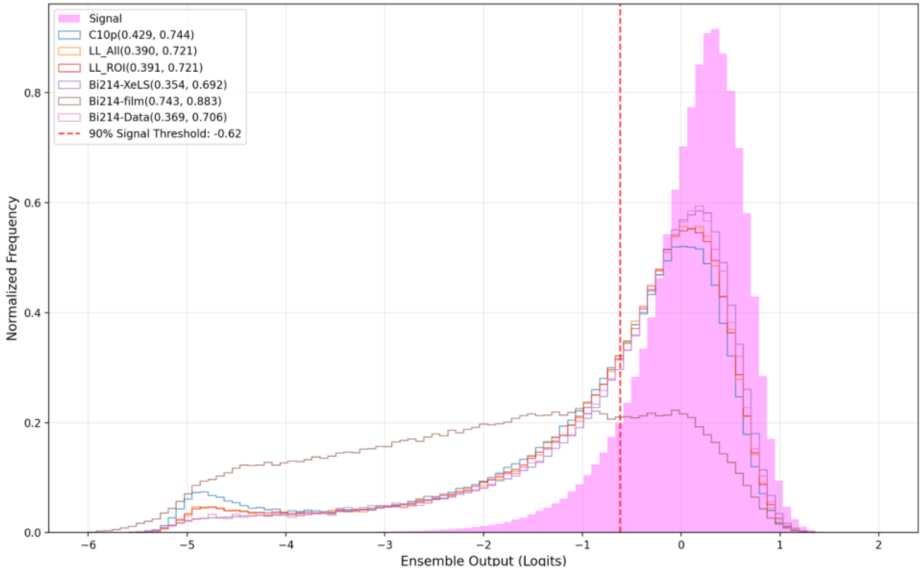
KamNet-ViViT (Parallel integration of temporal features)

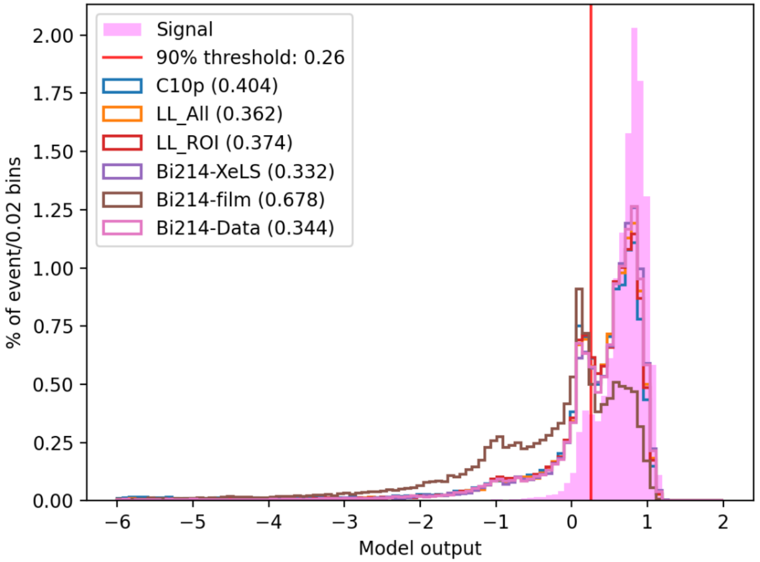
KamNet



ViViT







KamNet-ViViT

S2ViViT