

水チェレンコフ検出器における CNN 機械学習の手法を用いた
粒子識別の研究
(Study on Particle Identification Using CNN Machine Learning
Methods in Water Cherenkov Detector)

西嶋 遼太郎

2026 年 1 月 23 日

概要

スーパーカミオカンデは陽子崩壊、CP 対称性の破れ、超新星ニュートリノの観測などを目的とした大型水チェレンコフ検出器である。現在スーパーカミオカンデで用いられている再構成および粒子識別プログラムである fiTQun には計算コストが高いという課題がある。

そこで本研究ではより計算コストの低い粒子識別プログラムとして CNN を用いて、スーパーカミオカンデのシミュレーションデータに対して粒子識別を行なった。識別を行なった粒子は $e/\mu, e/\pi^0, e/\gamma$ の 3 組であり、各組について電子の運動量 $p_e = 200, 500, 1000$ MeV の 3 パターンを CNN のモデル MobileNetV3 を用いて学習を行った。そして学習の結果得られたモデルを用いて識別の正答率と計算時間を求めて fiTQun と比較した。また、各粒子および運動量について観測される光量の角度分布を比較することで、CNN による識別正答率の結果の説明を試みた。

得られた CNN による識別結果は e/μ での識別率は約 1.0 と高精度であった一方で e/π^0 では 0.7~0.9 で運動量が増加することによって精度が低下し、 e/γ の識別率は 0.5 と識別困難であった。この結果が e/μ では CNN が fiTQun を上回ったものの、 e/π^0 では CNN は fiTQun を上回ることができなかった。

CNN の計算時間については、1event あたり 0.02 秒であり、fiTQun の再構成時間の 10^{-4} と非常に短い時間での識別を達成した。

目次

第 1 章	序論	6
1.1	ニュートリノ	6
1.2	ニュートリノ振動	6
1.3	ニュートリノ実験	8
1.4	論文の概要	12
第 2 章	スーパーカミオカンデ	13
2.1	スーパーカミオカンデ検出器	13
2.2	検出原理	14
2.3	DAQ システム	15
2.4	SK のフェーズ	17
2.5	水循環・浄化システム	17
2.6	事象選別	17
2.7	事象再構成	20
2.8	粒子識別	22
第 3 章	畳み込みニューラルネットワーク	24
3.1	ニューラルネットワーク	24
3.2	畳み込みニューラルネットワーク	26
3.3	MobileNetV3	28
第 4 章	CNN による粒子識別の方法	31
4.1	物理背景	31
4.2	光量調整	31
4.3	シミュレーションデータ生成	32
4.4	データの前処理	33
4.5	ハイパーパラメータの調整	34
4.6	CNN による学習	35
第 5 章	CNN による粒子識別の結果	36
5.1	e と μ の識別	36
5.2	e と γ の識別	36
5.3	e と π^0 の識別	37

第 6 章	考察	40
6.1	fiTQun と CNN の性能比較	40
6.2	光量の角度分布	41
第 7 章	まとめ	43
参考文献		44

目次

1.1	pp-chain	8
1.2	SNO による太陽ニュートリノの測定結果 横軸： ν_e フラックス 縦軸： ν_μ, ν_τ フラックス 点線は標準太陽モデルによる予想値を示す。[1]	9
1.3	SK による e -like イベントと μ -like イベントの天頂各分布 斜線の領域はニュートリノ振動がないとした場合の MC による予想値で、実線はニュートリノ振動があるとしてデータに fit した値を示す [2]	10
1.4	左図：T2K で観測された ν_e 候補と $\sin^2 2\theta_{13} = 0.144$ で fit した MC による予想値 右図： δ_{cp} の関数として $\sin^2 2\theta_{13}$ の 68%, 90%CL。実線は $\sin^2 2\theta_{13}$ の最良値を示す。上図はニュートリノ質量階層について正常階層、下図は逆階層と想定している。[3]	11
2.1	SK タンクのスケッチ [5]	13
2.2	SK の PMT[5]	14
2.3	チェレンコフリング	15
2.4	QTC とその周り [6]	15
2.5	QTC のタイミングチャート	16
2.6	QTC の仕様	16
2.7	SK の水清浄化システム [8]	17
2.8	FC 1-リング Sub-GeV(左図), Multi-GeV(右図) イベントの PID 尤度関数 ($\log(L_e/L_\mu)$) による e -like イベントと μ -like イベントの識別。点が観測データ、ヒストグラムが大気ニュートリノ MC を表している [9]	22
2.9	e イベント (左) と π^0 イベント (右) 縦軸：対数尤度比 横軸： π^0 fit によって再構成された不変質量 マゼンタの線が e と π^0 を識別する線	23
3.1	簡単なニューラルネットワークの例	24
3.2	ニューロン間の入出力	25
3.3	畳み込み演算の例	27
3.4	パディングの例	27
3.5	MaX プーリングの例	28
3.6	MobileNetV3 の仕様 Operator の Bneck は Bottle neck の略。SE は Squeeze and Excitation モジュールをそのブロックで使用するかどうかを表す。NL は活性化関数を表し、HS は h-swish 関数、RE は ReLU 関数を意味する。	29

3.7	ボトルネック構造 [10]	29
3.8	squeeze and excitation モジュール	30
4.1	各粒子の運動量と総光量の関係 左上: e 、右上: μ 、左下 π^0	32
4.2	SKdetsim で生成したイベントの例。左上が e イベント、右上が μ イベント、左下が γ イベント、右下が π_0 イベントに対応する。4つの粒子ともバーテックスは中心、粒子の運動量方向は斜め下にした	33
4.3	上底面を長方形に変換する方法	34
4.4	円筒画像から正方形画像への変換	34
4.5	ハイパーパラメータ調整 各モデルの accuracy	35
5.1	運動量 200 MeV/c の e と運動量 396 MeV/c の μ の識別における accuracy	37
5.2	運動量 500 MeV/c の e と運動量 685 MeV/c の μ の識別における accuracy	37
5.3	運動量 1000 MeV/c の e と運動量 1154 MeV/c の μ の識別における accuracy	37
5.4	運動量 200 MeV/c の e と運動量 200 MeV/c の γ の識別における accuracy	38
5.5	運動量 500 MeV/c の e と運動量 500 MeV/c の γ の識別における accuracy	38
5.6	運動量 1000 MeV/c の e と運動量 1000 MeV/c の γ の識別における accuracy	38
5.7	運動量 200 MeV/c の e と運動量 146 MeV/c の π^0 の識別における accuracy	39
5.8	運動量 500 MeV/c の e と運動量 480 MeV/c の π^0 の識別における accuracy	39
5.9	運動量 1000 MeV/c の e と運動量 991 MeV/c の π^0 の識別における accuracy	39
6.1	e の光量角度分布 (青線) と μ, γ, π^0 の光量角度分布 (赤線) 左上: e のみ、右上: e と μ 、左下: e と γ 、右下: e と π^0 運動量は $(p_e, p_\mu, p_\gamma, p_{\pi^0}) = (500 \text{ MeV/c } 685 \text{ MeV/c } 500 \text{ MeV/c } 480 \text{ MeV/c})$	42
6.2	e の光量角度分布 (青線) と π^0 の光量角度分布 (赤線) 左上: $p_e = 200 \text{ MeV/c}$ 右上: $p_e = 500 \text{ MeV/c}$ 左下: $p_e = 1000 \text{ MeV/c}$	42

表目次

1.1	R-ratio の SK と IMB による測定値	10
2.1	SK の各フェーズにおける状態	17
2.2	fiTQun による e/μ の誤識別率	22
2.3	fiTQun により再構成および識別にかかった時間	22
2.4	fiTQun による e/π^0 の誤識別率	23
4.1	各粒子種の平均総光量が等しくなるような運動量	32
4.2	生成した各粒子の運動量	33
4.3	ハイパーパラメータの調整のための各モデル Lr は learning rate Wd は weight decay の略	34
4.4	CNN のモデルとハイパーパラメータ	35
5.1	e と μ の運動量のペアと 100 epoch でのモデルによる e/μ の識別正答率	36
5.2	e と γ の運動量のペアと 100 epoch でのモデルによる e/γ の識別正答率	36
5.3	e と π^0 の運動量のペアと 100 epoch でのモデルによる e/π^0 の識別正答率	38
5.4	e, π^0 それぞれの accuracy $p_e = 500 \text{ MeV/c}$ での 96~100 epoch での平均値	38
6.1	CNN と fiTQun による粒子識別の accuracy の比較	40
6.2	CNN による計算時間	41
6.3	CNN と fiTQun による粒子識別に要する時間の比較	41

第 1 章

序論

1.1 ニュートリノ

ニュートリノは標準理論においてレプトンに分類され、電荷 0、スピン 1/2 の素粒子である。ニュートリノには 3 世代があり、電子ニュートリノ、ミューオンニュートリノ、タウニュートリノと呼ばれる。これらは同じ世代のレプトン、電子、ミューオン、タウオンと、荷電カレント (CC) 相互作用によって入れ替わる。一方で、中性カレント (NC) 弱相互作用ではニュートリノは種類を変えずに相互作用する。ニュートリノは弱い相互作用のみをするため、ほとんど物質と反応しない。

1.2 ニュートリノ振動

ニュートリノ振動とは、ニュートリノが飛行中にフレーバーを変えることである。この現象の発見により、ニュートリノの質量が 0 ではないことがわかり、標準模型を超える物理の発見となった。この節ではニュートリノ振動の基礎的な理論を説明する。

1.2.1 3 世代のニュートリノ振動

まず、ニュートリノが正規直行する固有状態で書けるとし、各フレーバー固有状態はユニタリーな混合行列 U と質量固有状態によって表すことができるとする。

$$\begin{aligned} |\nu_\alpha\rangle &= \sum_i U_{\alpha i} |\nu_i\rangle & |\nu_i\rangle &= \sum_\alpha U_{\alpha i}^* |\nu_\alpha\rangle \\ U^\dagger U &= 1 & \sum_i U_{\alpha i} U_{\beta i}^* &= \delta_{\alpha\beta} & \sum_\alpha U_{\alpha i} U_{\alpha j}^* &= \delta_{ij} \end{aligned}$$

質量固有状態 ν_i は定常状態で、次の時間依存性を示す。

$$|\nu_i(x, t)\rangle = e^{-iE_i t} |\nu_i(x, 0)\rangle$$

ニュートリノが $x = 0$ で運動量 p で放出されたとすると、

$$|\nu_i(x, 0)\rangle = e^{ipx} |\nu_i\rangle$$

相対論的な時、

$$E_i = \sqrt{m_i^2 + p_i^2} \simeq p_i + \frac{m_i^2}{2p_i} \simeq E + \frac{m_i^2}{2E}$$

フレーバーニュートリノ $|\nu_\alpha\rangle$ が $t=0$ に運動量 p で放出されたとすると、

$$|\nu(x, t)\rangle = \sum_i U_{\alpha i} e^{-iE_i t} |\nu_i\rangle = \sum_{i, \beta} U_{\alpha i} U_{\beta i}^* e^{ipx} e^{-iE t} |\nu_\beta\rangle$$

よってフレーバー変換 $\nu_\alpha \rightarrow \nu_\beta$ の振幅の時間依存性は

$$A(\alpha \rightarrow \beta)(t) = \langle \nu_\beta | \nu(x, t) \rangle = \sum_i U_{\beta i}^* U_{\alpha i} e^{ipx} e^{-iE t}$$

距離の関数として表すと、

$$A(\alpha \rightarrow \beta)(t) = \sum_i U_{\beta i}^* U_{\alpha i} \exp(-i \frac{m_i^2}{2} \frac{L}{E}) = A(\alpha \rightarrow \beta)(L)$$

遷移確率 P は遷移振幅 A から得られて、

$$\begin{aligned} P(\alpha \rightarrow \beta)(t) &= |A(\alpha \rightarrow \beta)|^2 = \sum_i \sum_j U_{\alpha i} U_{\alpha j}^* U_{\beta i}^* U_{\beta j} e^{-i(E_i - E_j)t} \\ &= \sum_i |U_{\alpha i} U_{\beta i}|^2 + 2\text{Re} \sum_{j>i} U_{\alpha i} U_{\alpha j}^* U_{\beta i}^* U_{\beta j} \exp(-i \frac{\Delta m^2}{2}) \frac{L}{E} \end{aligned}$$

1.2.2 2 フレーバーでのニュートリノ振動

ここでは ν_e, ν_μ 間の振動を考える。 ν_e, ν_μ と ν_1, ν_2 の関係が次のように表せる。

$$\begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

前節の式を用いて対応する遷移確率は、

$$\begin{aligned} P(\nu_e \rightarrow \nu_\mu) &= P(\nu_\mu \rightarrow \nu_e) \\ &= \sin^2 2\theta \times \sin^2 \left(\frac{\Delta m^2}{4} \frac{L}{E} \right) \\ &= \sin^2 2\theta \times \sin^2 \left(\pi \frac{L}{L_0} \right) \\ \text{with } L_0 &= 2.48 \frac{E/\text{MeV}}{\Delta m^2/\text{eV}^2} m \end{aligned}$$

1.2.3 ニュートリノ振動における CP violation

ニュートリノ振動における CP violation とは $\nu_\alpha \rightarrow \nu_\beta$ の振動確率が $\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta$ の振動確率と異なることを意味する。すなわち、

$$\Delta P_{\alpha\beta}^{CP} = P(\nu_\alpha \rightarrow \nu_\beta) - P(\bar{\nu}_\alpha \rightarrow \bar{\nu}_\beta) \neq 0 \quad (\alpha \neq \beta)$$

PMNS 行列

$$U_{PMNS} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & C_{23} & S_{23} \\ 0 & -S_{23} & C_{23} \end{pmatrix} \begin{pmatrix} C_{13} & 0 & S_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -S_{13}e^{-i\delta} & 0 & C_{13} \end{pmatrix} \begin{pmatrix} C_{12} & S_{12} & 0 \\ -S_{12} & C_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

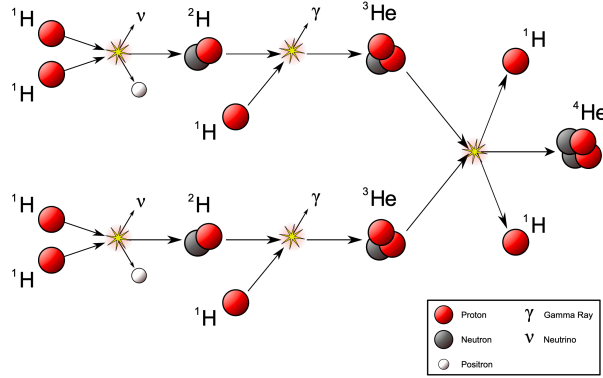


図1.1 pp-chain

を用いると、

$$\Delta P_{\alpha\beta}^{CP} = -16J_{\alpha\beta}\sin\left(\frac{\Delta m_{12}^2}{4E}L\right)\sin\left(\frac{\Delta m_{23}^2}{4E}L\right)\sin\left(\frac{\Delta m_{13}^2}{4E}L\right)$$

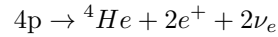
$$J_{\alpha\beta} \equiv \text{Im}[U_{\alpha 1}U_{\alpha 2}^*U_{\beta 1}^*U_{\beta 2}] = \pm c_{12}s_{12}c_{23}s_{23}c_{13}^2s_{13}\sin\delta$$

と表すことができる。よって全ての混合角が $0, \frac{\pi}{2}, \pi$ でなく、 $\sin\delta \neq 0$ を示すことができれば CP が破れていると言える。

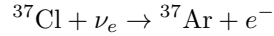
1.3 ニュートリノ実験

1.3.1 太陽ニュートリノ

太陽では核融合によりエネルギーが作られている。水素が融合してヘリウムを生み出す過程は主に次のようである。



初の太陽ニュートリノ実験は 1968 年アメリカで行われたホームステーキ実験で、ニュートリノを検出した反応は



であり、エネルギー閾値は 814 keV であった。標準太陽モデルによって予想された値は 7.5 ± 1.0 SNU であり、主に ${}^8\text{B}$ ニュートリノから構成されている。しかし、実際に観測された値は $2.56 \pm 0.16(\text{stat.}) \pm 0.15(\text{sys.})$ SNU と理論値の約 1/3 であった。その後、いくつか行われた太陽ニュートリノ実験でも観測された電子ニュートリノが、理論値よりも少ないことが確認され、この問題は太陽ニュートリノ問題と呼ばれていた。

この問題に対して、決定的となる実験結果を出したのが 1999 年から開始された SNO 実験である。SNO 実験では 1000 t の重水を使ったチェレンコフ検出器による観測が行われた。この検出器でのニュートリノの反応

は複数あり、それは

$$\nu_e + d \rightarrow e^- + p + p \quad (\text{CC})$$

$$\nu + e^- \rightarrow \nu + e^- \quad (\text{ES})$$

$$\nu + d \rightarrow \nu + p + n \quad (\text{NC})$$

の3つである。このうち、CCは ν_e のみに感度があり、ESはすべてのフレーバーに感度があるが、 ν_e が支配的である。一方で、NCはすべてのフレーバーに同等の感度があるため、全ニュートリノ数を決定できる。

SNO 実験の測定結果は図1.2のようであり、全ニュートリノフラックスは標準太陽モデルの予想と一致した

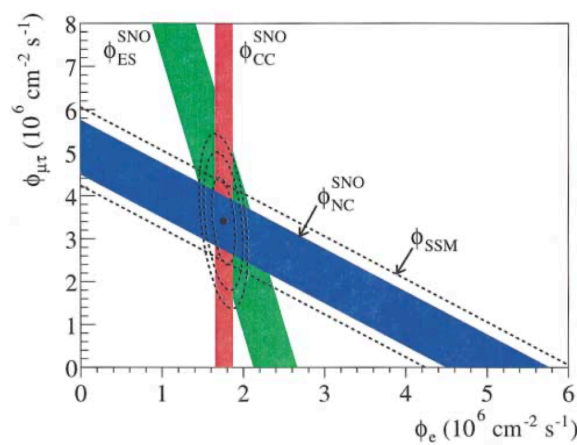


図1.2 SNO による太陽ニュートリノの測定結果 横軸： ν_e フラックス 縦軸： ν_μ, ν_τ フラックス 点線は標準太陽モデルによる予想値を示す。[1]

のに対し、電子ニュートリノフラックスは全ニュートリノフラックスの約 1/3 であった。この測定結果により、ニュートリノは失われているのではなく、他のフレーバーに変化していることが示唆された。

1.3.2 大気ニュートリノ

大気ニュートリノは、宇宙線と大気の相互作用によって生じたパイオンやミューオンの崩壊によって生まれるニュートリノのことである。

$$\pi^+ \rightarrow \mu^+ \nu_\mu \quad \mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu$$

$$\pi^- \rightarrow \mu^- \bar{\nu}_\mu \quad \mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$$

ニュートリノ振動の指標として次の式で定義される R -ratio が重要になる。

$$R = \frac{[N(\mu - \text{like})/N(e - \text{like})]_{\text{obs}}}{[N(\mu - \text{like})/N(e - \text{like})]_{\text{exp}}}$$

大気ニュートリノの観測はスーパーカミオカンデ (SK) 実験や IMB などが行っていて、 R -ratio の測定値は表1.1のようであった。

Experiment	R
SK(sub-GeV)	$0.638 \pm 0.017 \pm 0.050$
SK(multi-GeV)	$0.675 \pm 0.034 \pm 0.080$
IMB	$0.54 \pm 0.05 \pm 0.11$

表1.1 R-ratio の SK と IMB による測定値

表1.1からわかるように R -ratio はおよそ 0.6 と 1 からずれていて、これはニュートリノ振動が起きていることを示唆している。さらに詳しいことはニュートリノの頂点角分布を見ることでわかる。

SK による ν_e と ν_μ の頂点角分布は図1.3に示されている。図1.3からわかるように、MC の予想値に一致して

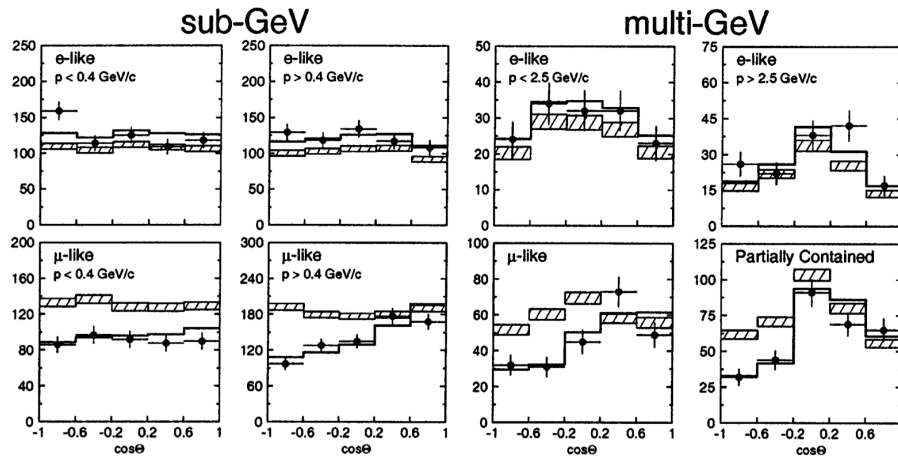


図1.3 SK による e -like イベントと μ -like イベントの天頂角分布

斜線の領域はニュートリノ振動がないとした場合の MC による予想値で、実線はニュートリノ振動があるとしてデータに fit した値を示す [2]

いる e -like イベントと異なり、 $\cos\theta$ が小さくなるほど、予想値と観測値が離れていっている。この結果から $\nu_\mu \rightarrow \nu_\tau$ のニュートリノ振動が起きていると説明できる。

1.3.3 超新星ニュートリノ

超新星は2つの異なる星の終状態から発生する。一つは連星系が融合してチャンドラセカル質量を超えたとき、もう一つは巨星内部でこれ以上核融合によってエネルギーを生み出せなくなり、重力的に不安定になったときである。

重力崩壊が起こっている大質量性のコアでは高温高密度のため電子捕獲反応 $p + e \rightarrow n + \nu_e$ が起こる。この時、星内部には密度が 10^{11} g/cm^3 以上と非常に高いためにニュートリノが星の外に出られない領域があり、電子ニュートリノの閉じ込めが起こる。この閉じ込められたニュートリノが超新星爆発時に一斉に放出されルコとで超新星ニュートリノとして検出される。

超新星ニュートリノの検出するためのリアルタイム処理する検出器としては、SK などのチェレンコフ検出器か KamLAND などの液体シンチレータが挙げられる。

実際に観測された超新星ニュートリノイベントは SN1987A であり、カミオカカンデでは 12 イベント、IMB では 8 イベントが観測された。

もう少し書く

1.3.4 加速器ニュートリノ

加速器ニュートリノ実験では、加速器側で ν_μ ビームを生み出し、 $O(100\text{ km})$ だけ離れた検出器でビームを検出する。 ν_μ ビームを生成する方法は、加速した陽子ビームを炭素などの標的に衝突させて荷電パイオンを生成させ、電磁ホーンによってパイオンの方向を制御する。飛行中にパイオンが μ と ν_μ に崩壊し、ミューオンは壁で止められる。残った ν_μ が前置検出器および遠方検出器で検出される。

代表的な加速器ニュートリノ実験としては T2K 実験があり、J-PARC で生成した ν ビームを SK で検出する実験である。

ν ビームは 30 GeV の陽子から生成され、炭素標的から 280 m の位置にある前置検出器でニュートリノ振動前の ν ビームの性質を計測し、295 km の位置にある SK で振動したニュートリノを観測する。図1.4の左図は T2K 実験で観測された ν_e 候補イベントについての図であり、右図は観測結果から計算された $\sin^2 2\theta_{13}$ の値を示している。この実験によって初めて $\nu_\mu \rightarrow \nu_e$ が観測された。

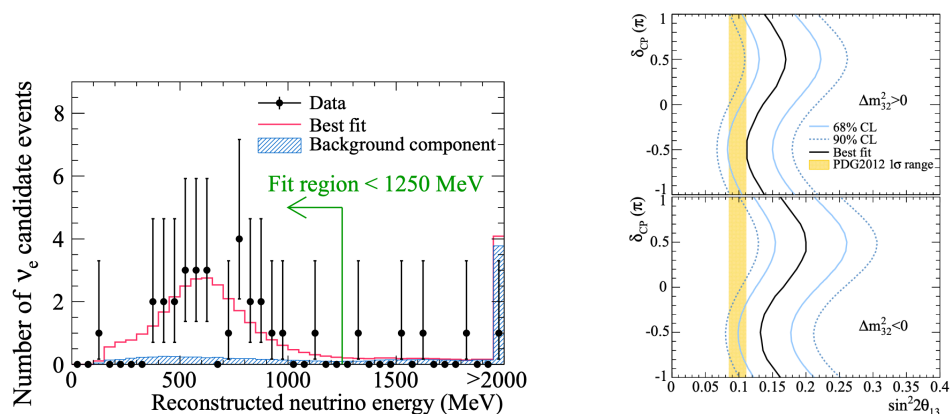


図1.4 左図：T2Kで観測された ν_e 候補と $\sin^2 2\theta_{13} = 0.144$ で fit した MC による予想値
右図： δ_{CP} の関数として $\sin^2 2\theta_{13}$ の 68%, 90% CL。実線は $\sin^2 2\theta_{13}$ の最良値を示す。上図はニュートリノ質量階層について正常階層、下図は逆階層と想定している。[3]

1.3.5 ハイパーカミオカンデ

ハイパーカミオカンデはカミオカンデ、スーパーカミオカンデに続く次世代型大型水チェレンコフ検出器で、岐阜県神岡町に建設中である。大きさは直径 68 m、高さ 71 m、水容量 26 万トンであり、2028 年に実験が開始される予定である。主な研究目的としては、ニュートリノの CP 対称性の破れの発見、陽子崩壊の探索、超新星爆発からのニュートリノの観測があげられる。

1.4 論文の概要

続く2章ではスーパーカミオカンデについて、特に粒子識別を行う `fiTQun` に注目して説明を行う。3章では畳み込みニューラルネットワークについて、ニューラルネットワークとは何かから説明し、また最後の節では本研究で扱ったモデルである `MobileNetV3` について簡単に記述している。4章では本研究で行なったCNNによる粒子識別の方法について説明し、5章で結果を6章で結果に対する考察を行なう。そして、最後の7章で本研究の結論を記述している。

第 2 章

スーパーカミオカンデ

スーパーカミオカンデとは岐阜県神岡町に位置する巨大水チェレンコフ検出器である (図2.1)。その大きさは直径 39 m、高さ 42 m の円柱形であり、水容量は 50,000 トンに及ぶ。スーパーカミオカンデ実験の目的は、陽子崩壊、太陽ニュートリノ、大気ニュートリノ、超新星ニュートリノ、T2K 実験などの人工ニュートリノの観測が挙げられる。

2.1 スーパーカミオカンデ検出器

SK の水槽は池の山地下 1000 m に位置し、水容量 50,000 トンのタンク内は内部検出器と外部検出機に光学的に分かれている。内部検出機には 20 inch PMT が 11,129 個、外部検出機には 8 inch PMT が 1,885 個設置されており、内部検出器が主検出器として、外部検出器は veto 検出器として機能している。SK タンクが地下 1000 m にあることで、バックグラウンドとなる宇宙線ミューオンが 1000 m の岩盤によって遮蔽され、SK タンクに到達する数を大幅に減じている。また、PMT は磁場による影響を受けやすいため、磁気補償コイルをタンクの壁に設置することで地磁気を相殺して、残存磁場を 50 mG 以下にまで減らしている。

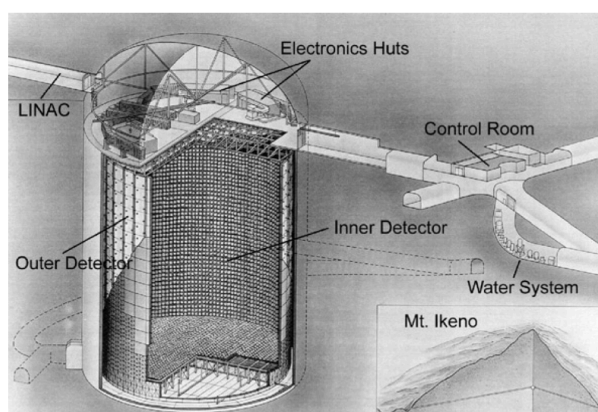


図2.1 SK タンクのスケッチ [5]

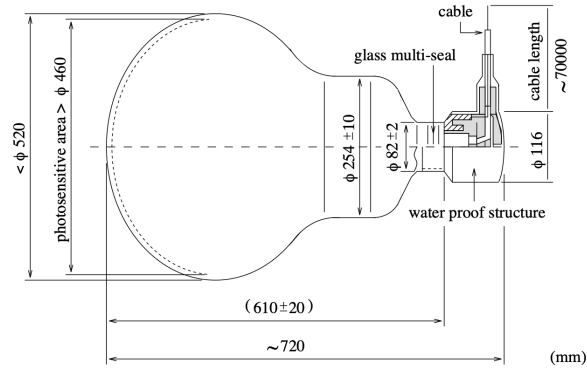


図2.2 SK の PMT[5]

2.1.1 内部検出器 (ID)

SK タンクは PMT を設置する円柱形の構造体によって 2 つの領域に分けられていて、内側は直径 33.8 m、高さ 36.2 m で 32,000 トンの水が内包され、これが内部検出器 (ID) と呼ばれる。ID の PMT は 70 cm 間隔で 11,129 個配置されていて、ID 表面のうち 40 % が光電面によってカバーされている。20 inch PMT(図2.2) はバイアルカリ (Sb-K-Cs) 光電面でピーク QE は 21 % (360 ~ 400 nm) である。ID PMTs にかけている電圧は 1700 ~ 2000 V であり、ゲイン 10^7 がとなるように設定されている。

2.1.2 外部検出器 (OD)

分けられた領域の外側を外部検出器 (OD) とよび、1,885 個の PMT が設置されている。OD PMT は ID よりも疎であるため、波長変換板を用いることと、tyvek sheet を OD 表面につくることによって光収集効率をあげている。OD は宇宙線ミューオンや周囲の岩盤からの放射線によるバックグラウンドの veto 検出器として機能している。

2.2 検出原理

SK は水チェレンコフ検出器という名前の通り、チェレンコフ光を PMT で検出することでタンクに入射する粒子を検出している。チェレンコフ光とは荷電粒子が屈折率 n の媒質中を光速 c/n を超える速度で運動するときに放射される光である。この光は円錐状に伝播していき (図2.3)、円錐の開口角 θ は

$$\cos \theta = \frac{1}{n\beta}$$

また、荷電粒子が単位長さ進むあたり、単位波長あたりの光子数は

$$\frac{d^2 N}{dx d\lambda} = \frac{2\pi\alpha}{\lambda^2} \left(1 - \frac{1}{n^2\beta^2}\right)$$

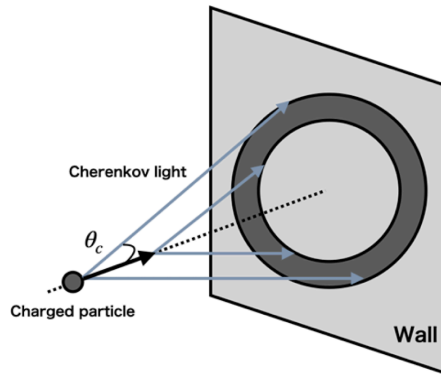


図2.3 チェレンコフリング

媒質が水の場合、屈折率 n が $n = 1.33$ なので開口角 θ は約 42° である。また、このときチェレンコフ閾値は電子について $0.569 \text{ MeV}/c$ 、ミューオンについて $115.7 \text{ MeV}/c$ 、陽子について $1.04 \text{ GeV}/c$ である。

2.3 DAQ システム

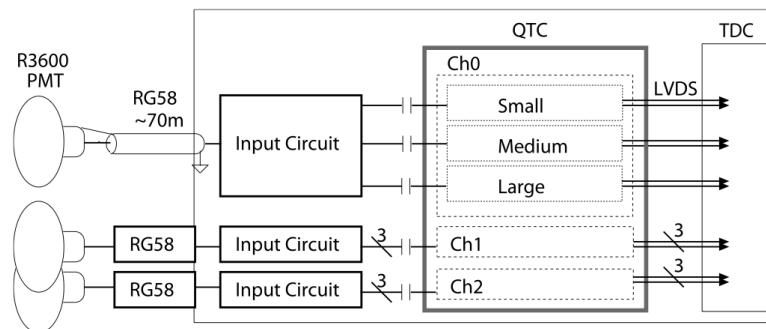


図2.4 QTC とその周り [6]

図2.4はPMTからの信号を処理するQTC（Charge to Time converter）とその周りの機器を表している。チェレンコフ光がPMTの光電面に入射すると光電効果で電子が放出され、放出された電子はPMTの増幅部で約 10^7 倍に増幅された信号として出力される。PMTの出力信号は同軸ケーブルを通過して、3つの信号に分けられてQTCに入力される。QTCの3つの入力是小中大の3つのゲインに対応し、これによって電荷についての広いダイナミックレンジと十分な分解能を確保している。QTCに入った信号がディスクリミネータの閾値（ $\frac{1}{4} \text{ p.e.}$ ）を超えると電荷の積分が始まる。また、同時にhitの信号が出される。電荷の積分は 400 ns だけ行われ、その後一定電流での放電が行われる。出力信号は電荷積分が始まった時から、積分電荷が放電によってコンパレータの閾値を下回ったときまでの間、一定電圧で出力される。放電によってコンパレータの閾値を下回るまでの時間は積分電荷の大きさによって変わるため、出力信号の立ち下がり時間から積分電荷を、立ち上がり時間からhitの時間を読み取ることができる。図2.5はQTCのタイミングチャートを表している。ま

た、図2.6には QTC の仕様を示してある。

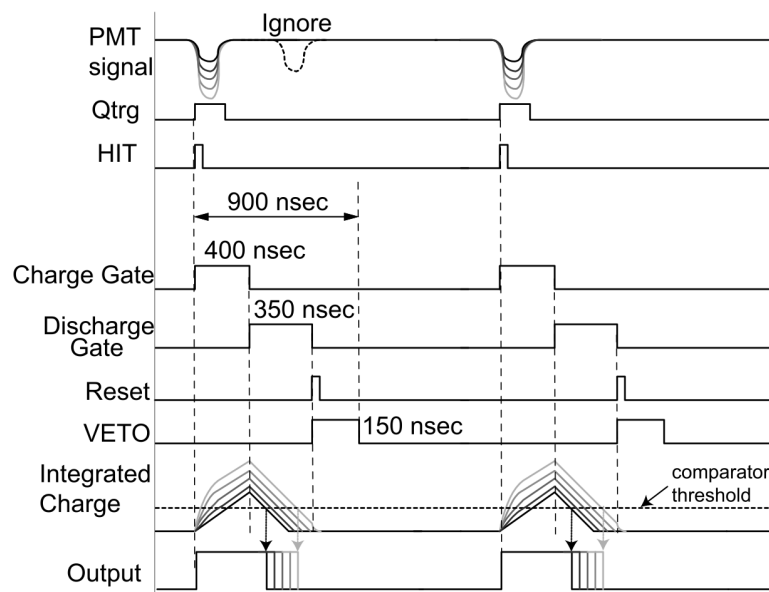


図2.5 QTC のタイミングチャート

Type of trigger	self trigger by discriminator
Number of Input Channels	3
Processing Speed	~ 900 ns/cycle
Charge Integration Gate	400 ns
Number of Gains	3 (Ratio 1 : $\frac{1}{7}$: $\frac{1}{49}$)
Discriminator Threshold	-0.3 ~ -14 mV (small range)
Charge Dynamic Range	0.2 ~ 51 pC (small) 1 ~ 357 pC (medium) 5 ~ 2500 pC (large)
Charge Resolution	~ 0.2 pC (small)
Integral (Non-)Linearity	< $\pm 1\%$
Timing Resolution	0.3 ns (2 pC, -3 mV) < 0.1 ns (> 100 pC)
Power Dissipation	< 100 mW/ch
Process	0.35 μm CMOS
Package	100 pin CQFP

図2.6 QTC の仕様

QTC の出力は TDC (Time to Digital converter) に入力され、TDC では入力の立ち上がりとしち下がりを出して信号の幅を測定する。その後 TDC からの出力は FPGA(Field Programmable Gate Array) に渡され、電荷、タイミング、入力チャンネルの情報を含んだ 6 バイトセルに変換され、イーサネット でフロントエンド PC に送信される。

2.4 SK のフェーズ

SK の実験は 1996 年から始まり、今に至るまで複数回のアップデートを経てきた。この節では SK-I から SK-VIII までの各フェーズについてまとめる

フェーズ	運転期間	IDPMT の本数	エレクトロニクス	Gd 濃度	備考
SK-I	1996.4~2001.7	11,146	ATM	0 %	PMT の破損事故 (2001)
SK-II	2002.12~2005.10	5,182	ATM	0 %	PMT に防爆カバーを装着
SK-III	2006~2008.8	11,129	ATM	0 %	
SK-IV	2008.9~2018.6	11,129	QBEE	0 %	
SK-V	2019.1~2020.6	11,129	QBEE	0 %	SK-Gd の準備フェーズ
SK-VI	2020.7~2022.6	11,129	QBEE	0.011 %	
SK-VII	2022.6~2024.9	11,129	QBEE	0.03 %	補償コイルの一部が故障 (2023)
SK-VIII	2024.9~	11,129	QBEE	0.033 %	補償コイル修理完了

表2.1 SK の各フェーズにおける状態

2.5 水循環・浄化システム

チェレンコフ光の透過性や自然放射線を抑える観点からタンク内の水は清浄である必要があるため、SK では水清浄化システムを用いている。図2.7は SK-IV の水清浄化システムを表した図である。このシステムによってタンク内の水は 60 ton/h で循環され、水温は約 13 度に保たれている。

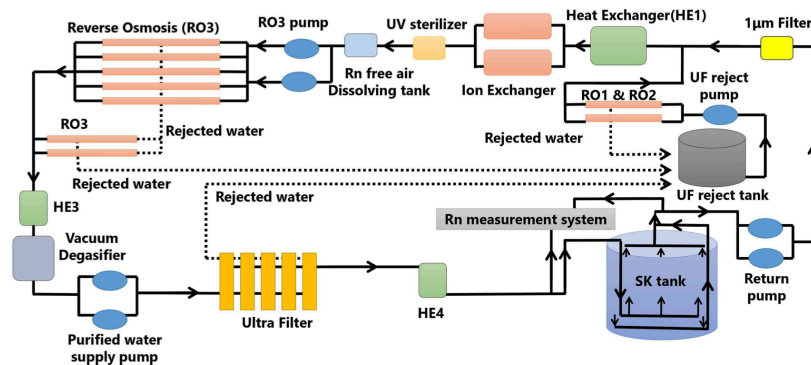


図2.7 SK の水清浄化システム [8]

2.6 事象選別

オフラインプロセスでの事象選別

DAQ からの出力はオンライントリガーをかけられた後、オフライン校正をされる。この段階のイベントレートは約 $10^6/\text{day}$ であり、宇宙線ミューオンイベントや水中に含まれるラドンなどの放射性物質によるイベ

ントが多く含まれている。事象選別ではこれらのバックグラウンド事象を選別して取り除き、より効率的に物理解析を行う。

SK で観測される大気ニュートリノイベントは大きく 3 つに分類される。すなわち、Fully contained(FC),Partially contained(PC),Upward-going muon(Up-mu) に分けられる。FC は粒子が内部検出器内でのみエネルギーを落とし、外部検出器に何の反応もないイベント。PC は内部検出器の外でもエネルギーを落としているイベント。Up-mu は ID の周りでニュートリノ反応によって生成されたミューオンが上向きに ID 中を移動したイベント。

本研究では FC イベントのみを扱うので、この節では FC イベントの事象選別のみを記述する。FC イベントの事象選別は 5 つのステップからなる。

2.6.1 1 つ目の選別

1 つ目の選別は以下の 2 つからなる。

1. ID PMT の総電荷が 200 p.e. 以下のイベントを除去する
2. イベントトリガー周りの 800 ns で OD PMT の hit 数が 55 よりも大きいイベントを除去する

これらは宇宙線ミューオンによるイベントと水に含まれる不純物による低エネルギーの放射線イベントを取り除くことを目的としている。これらのカットによってイベントレートは約 $10^3/\text{day}$ にまで減る。

2.6.2 2 つ目の選別

2 つ目の選別は以下の 2 つからなる

1. OD PMT の hit が 30 よりも大きくかつ、ID PMT の総電荷が 100,000 p.e. よりも小さいイベントを除去する
2. 1 つの ID PMT の電荷が ID PMT の総電荷の半分を超えているイベントを除去する

1 つ目のカットはより低エネルギーの宇宙線ミューオンイベントをカットしつつ、超高エネルギーの FC イベントで OD に光漏れが起きているイベントを保持している。2 つ目のカットはいずれか一つの PMT での電氣的ノイズイベントを取り除いている。これらのカットによって、イベントレートは約 $10^2/\text{day}$ となる。

2.6.3 3 つ目の選別

3 つ目の選別は 1,2 つ目の選別で除去できなかったミューオンイベントやノイズイベントを除去することを目的とする以下の複数のカットからなる。

ハードミューオン

1 TeV 以上の高エネルギーミューオンはハードミューオンと呼ばれ、次のカットで除去される。

1. メイントリガー後の 500 ns 間での OD hit が 40 以上のイベントを除去する

貫通ミューオン

貫通ミューオンについてのフィットは次の条件を満たすときにかけられる。

1. 1 つの ID PMT によって観測された電子が 231 p.e. よりも大きい
2. ID PMT の hit 数が 1000 以上

フィットの適合度が 0.75 以上の場合貫通ミューオンとして除去される。

停止ミューオン

貫通ミューオンのようにフィットをかけられ、適合度が 0 以上の時停止ミューオンとして除去される。

ケーブル穴ミューオン

SK にはケーブルを通すための穴が 12 個あり、そこには OD PMT がない。そのため、その穴を通ったミューオンは OD を hit させることなく ID に侵入できる。このミューオンを除去するために、ケーブル穴にはプラスチックシンチレータが設置され、シンチレータに反応が会った時、ケーブル穴ミューオンとして除去される。

フラッシャーイベント

フラッシャーイベントは PMT での放電によって引き起こされるイベントである。フラッシャーイベントは hit のタイミング分布が広がっている傾向にあるため、イベントトリガー後の 500 ns 間、100ns ほどの ID hit が常に 20 以上のイベントがフラッシャイベントとして除去される。

3 つ目の選別によってイベントレートは 45/day となる。

2.6.4 4 つ目の選別

4 つ目の選別では、残ったフラッシャーイベントをパターンマッチングによって除去する。フラッシャーイベントは同様の PMT hit パターンを繰り返す傾向にあるため、異なるイベント間での電荷パターンの相関を計算して、高い相関にあるイベントを除去する。この選別によってイベントレートは 18/day となる。

2.6.5 5 つ目の選別

5 つ目の選別は残ったミューオンイベントとフラッシャーイベントを除去することを目的とする。チェレンコフ閾値以下のミューオンは ID PMT に hit を生み出さないが、崩壊によって生じた電子によってイベントトリガーがかかることがある。このミューオンを取り除くために次のカットをかける。

1. トリガー前の -8800 ns から -100 ns の間のいずれかの 200 ns 間に OD hit が 9 以上のイベントを除去する

フラッシャーイベントについてより厳しいカットをかけるために、再構成によって得られた hit 時間をもとに 3 つ目の選別と同様のカットをかける。

5 つ目の選別によってイベントレートは約 16/day となる。

2.6.6 FC 選別のまとめ

最終的な FC ニュートリノイベントは次の条件を満たすイベント。

1. 再構成されたバーテックスが有効体積内にあること。ただし、有効体積とは ID 壁から 2 m 離れた内側の体積のこと

2. 最大の OD hit クラスターの hit 数が 16 より小さい

3. 可視エネルギーが 30 MeV 以上であること。

シミュレーションによると真の有効体積内のニュートリノイベントのうち、FC ニュートリノイベントと判定される確率は約 98 % であり、背景事象が混ざる確率は 0.1 % 以下である。また、最終的なイベントレートは～8/day となる。

2.7 事象再構成

SK で用いられている事象再構成法には、APfit, fiTQun, Bonsai がある。AP fit は現在大気ニュートリノの解析などには使われている。fiTQun は SK-IV で開発され、特にバーテックスの決定や粒子識別の精度において APfit よりも優れているが計算コストがかかる欠点がある。Bonsai は低エネルギー領域で用いられ、特に太陽ニュートリノの解析で役立っている。ここでは fiTQun の事象再構成法について詳しく説明した後、fiTQun を用いた粒子識別の精度や所要時間について議論する。

2.7.1 尤度関数

fiTQun では最尤法を用いている。SK のイベントの情報は各 PMT で hit したかどうか、hit した時間、検出された電荷である。ここで、粒子種やその運動量、バーテックスなどの情報すべてを含んだ仮説 \mathbf{x} を導入すると、尤度関数は次のようになる。

$$L(\mathbf{x}) = \prod_j^{\text{unhit}} P_j(\text{unhit}|\mathbf{x}) \prod_i^{\text{hit}} \{1 - P_i(\text{unhit}|\mathbf{x})\} f_q(q_i|\mathbf{x}) f_t(t_i|\mathbf{x})$$

この式において、 $P_j(\text{unhit}|\mathbf{x})$ は仮説 \mathbf{x} のに PMT_j が hit しなかった確率を表し、 $f_q(q_i|\mathbf{x})$ は仮説 \mathbf{x} において PMT_i の電荷が q_i になる確率を表し、 $f_t(t_i|\mathbf{x})$ は仮説 \mathbf{x} で PMT_i が hit した時間が t_i になる確率を表している。fiTQun では、この尤度関数を最大にするような仮説 \mathbf{x} を求めることで、粒子の種類、運動量、バーテックスを導き出している。

2.7.2 期待光電子数

実際の計算においては期待光電子数 μ 、すなわち仮説 \mathbf{x} のときの各 PMT における平均光電子数、を導入して次の尤度関数について考える。

$$L(\mathbf{x}) = \prod_j^{\text{unhit}} P_j(\text{unhit}|\mu_j) \prod_i^{\text{hit}} \{1 - P_i(\text{unhit}|\mu_i)\} f_q(q_i|\mu_i) f_t(t_i|\mathbf{x})$$

期待光電子数を計算するときには、直接 PMT に入射する光と散乱、反射して PMT に入射する光を別々に考えた上で最後に足し合わせる。直接光による期待光電子数を計算するには、粒子の軌跡上の各点からのチェレンコフ放射を各 PMT への距離、水の光透過率、PMT の光入射角依存性を考慮した上で、粒子軌跡上積分することで得られる。間接光による期待光電子数は、直接光の性質と事前にシミュレーションによって得られた散乱関数との積を粒子の軌跡上積分することで得られる。複数のチェレンコフリングがある場合には各リングについて個別に期待光電子数を計算したのちに足し合わせる。

2.7.3 unhit の確率

μ が平均光電子数なので、実際の光電子数は平均 μ の Poisson 分布に従う。そのため、unhit の確率は $e^{-\mu}$ となるはずである。しかし、実際には信号が一定の閾値を越えないと hit とならないため、3 次までの補正を含めて、次の式のように unhit の確率を近似している。

$$P(\text{unhit}|\mu) \approx (1 + a_1\mu + a_2\mu^2 + a_3\mu^3)e^{-\mu}$$

2.7.4 電荷尤度、時間尤度

電荷尤度は平均 μ_i の Poisson 分布と PMT で観測された電荷を比較することで得られる。

時間尤度は $f_t(t_i|t_i^{\text{exp}}, \Gamma, p, \mu_i)$ と表すことができる。ここで、 t_i^{exp} は予想される hit 時間、 Γ はいそう、 p は運動量を意味する。 t_i^{exp} は軌跡の midpoint で放出された光が直接 PMT に入射したときの時間を意味し、次の式で表される。

$$t_i^{\text{exp}} = t + s_{\text{mid}}/c + |\mathbf{R}_i^{\text{PMT}} - \mathbf{x} - s_{\text{mid}}\mathbf{d}|/c_n$$

ただし、 \mathbf{x} と t はバーテックスと粒子が作られた時間、 \mathbf{d} は粒子の方向 $\mathbf{R}_i^{\text{PMT}}$ は i 番目の PMT の位置、 s_{mid} は軌跡の長さの半分、 $c_n = c/n$ を意味する。時間尤度が期待光電子数に依存するのは、hit が記録されるのが一つ目の光子が PMT に到達したときであり、その分布の不定性は入射光子数が多いほど小さくなる。時間尤度の関数はパーティクルガンシミュレーションで決定される。直接光と間接光による時間尤度は別々に計算された後に合わせられる。また、複数リングがある場合も同様に個別のリングで計算した後に合わせられる。

2.7.5 バーテックスの prefit

尤度関数を最大化、実際には $-\ln L(x)$ を最小化する際に、局所最小値に止まってしまうことがないように、初期値を適切に決定することが重要になる。そこで fitQun では、時間の情報だけを用いて、vertex の位置と hit の時間を大まかに予測して初期値を決定する方法をとっている。予測は、以下の尤度関数を最大化する vertex の位置 \mathbf{x} と hit の時間 t を求めることで行なっている。

$$G(\mathbf{x}, t) := \sum_i^{\text{hit}} \exp(-(T_{\text{res}}^i/\sigma)^2/2)$$

$$T_{\text{res}}^i := t_i - t - |\mathbf{R}_{\text{PMT}}^i - \mathbf{x}|/c_n$$

パラメータ σ の値が prefit の精度を決定する。prefit は複数回行われ、段々パラメータ空間のグリッドの大きさと σ の値を小さくしていくことで効率的に高い精度で fit することができる。

2.7.6 1-リング再構成

バーテックスの prefit が終わったら、粒子の方向の初期値を決める。そのために、単位球上に等間隔に設定された 400 点について尤度を計算する。その次は運動量の初期値を決定するために、まず観測された総電荷から大まかな運動量の値を決め、その後運動量についての尤度スキャンをする。以上のように全てのパラメータの初期値を決定したら、全てのパラメータを同時に動かしながら MINIUT パッケージの SIMPLEX アルゴリズムによって $-\ln L(x)$ を最小化する。以上で最尤法によって予測された粒子のパラメータが決定される。

2.8 粒子識別

fiTQun による粒子識別は各粒子仮説ごとの best-fit 尤度関数を比較することによって行われる。例えば、 e と μ は $\log(L_e/L_\mu)$ の値によって識別される。図2.8が PID 尤度関数 $\log(L_e/L_\mu)$ による粒子識別の例を示している。

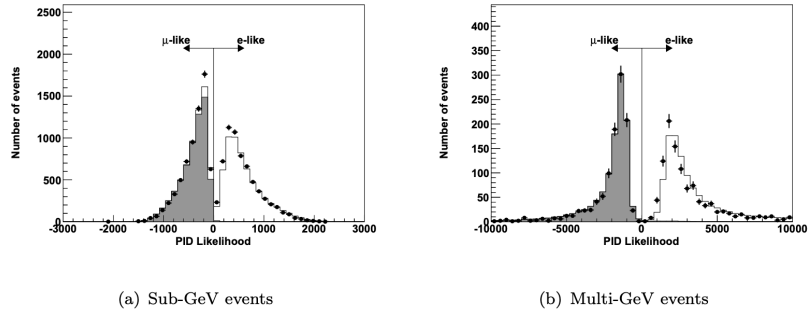


図2.8 FC 1ーリング Sub-GeV(左図),Multi-GeV(右図) イベントの PID 尤度関数 ($\log(L_e/L_\mu)$) による e -like イベントと μ -like イベントの識別。点が観測データ、ヒストグラムが大気ニュートリノ MC を表している [9]

2.8.1 fiTQun による e/μ 識別

fiTQun による e/μ の誤識別率は表2.3のようであった。

運動量 \ 粒子種	粒子種	
	e	μ
200 MeV/c	0.4 %	0.8 %
500 MeV/c	0.0 %	0.4 %
1000 MeV/c	0.0 %	0.1 %

表2.2 fiTQun による e/μ の誤識別率

また再構成および識別にかかった時間は表2.3のようであった。ただし、この時用いた CPU は Intel(R) Xeon(R) Gold 6348 CPU @ 2.60GHz であった。

運動量 (MeV/c)	時間 (s/events)
200	85.6
500	136.8
1000	177.9

表2.3 fiTQun により再構成および識別にかかった時間

2.8.2 fiTQun による e/π^0 識別

e と π^0 の場合も $\log(L_e/L_{\pi^0})$ の値によって識別するが、MC による e と π^0 イベントを $\pi^0\text{fit}$ によって再構成された不変質量と $\log(L_{\pi^0}/L_e)$ についてプロットすると図2.9のようになるため、識別には再構成された不変質量も用いて、

$$\log(L_{\pi^0}/L_e) < 175 - 0.875m_{\pi^0}(\text{MeV}/c^2)$$

を満たす場合に e と判定する。

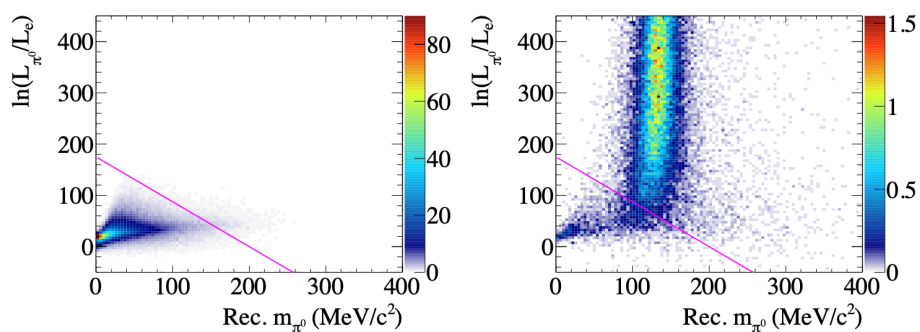


図2.9 e イベント (左) と π^0 イベント (右) 縦軸：対数尤度比 横軸： $\pi^0\text{fit}$ によって再構成された不変質量 マゼンタの線が e と π^0 を識別する線

fiTQun による e/π^0 の誤識別率は表2.4のようであった。

運動量	粒子種	
	e	π^0
200 MeV/c	5 %	2 %
500 MeV/c	12 %	5 %
1000 MeV/c	29 %	16 %

表2.4 fiTQun による e/π^0 の誤識別率

第 3 章

畳み込みニューラルネットワーク

本章では畳み込みニューラルネットワーク（CNN）について説明する。CNN は画像処理の分野で幅広く利用されており、本研究においても CNN の 1 モデルである MobileNetV3 を画像認識に使用している。

3.1 ニューラルネットワーク

ニューラルネットワークとは脳のニューロンの仕組みにヒントを得た機械学習の 1 種である。ニューラルネットワークは複雑なデータの関係性を捉えることができ、画像認識、音声認識、自然言語処理など多岐に渡って利用されている。

3.1.1 ニューラルネットワークの仕組み

ニューラルネットワークは入力層、一つ以上の隠れ層、出力層から構成されている。各層には複数のニューロンがあり、各ニューロンは前の層のすべてのニューロンからの出力を入力として受け取る。

ニューロン間のつながりは重みを持っていて、前の層からの出力に重みをかけた値とニューラルネットワー

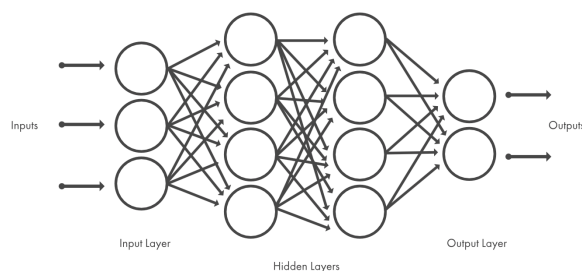


図3.1 簡単なニューラルネットワークの例

ク自身が持つバイアスとの総和を活性化関数によって変換した値が次のニューロンへの出力となる。活性化関数には sigmoid 関数 $f(x) = \frac{1}{1+e^{-x}}$ や ReLU 関数 $f(x) = \max(0, x)$ が用いられる。

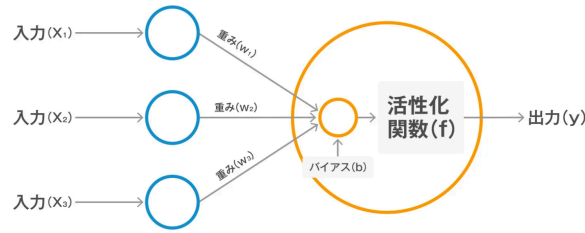


図3.2 ニューロン間の入出力

3.1.2 ニューラルネットワークの学習

ニューラルネットワークの学習とは訓練データから重みパラメータの値を決定することをいう。そのために用いられる指標は損失関数と呼ばれ、2乗和誤差 $E = \frac{1}{2} \sum_k (y_k - t_k)^2$ 、交差エントロピー誤差 $E = -\sum_k t_k \log y_k$ がよく使用される。ここで y_k はニューラルネットワークの出力、 t_k は教師データ、 k はデータの次元数を表す。この損失関数が小さくなるように重みパラメータを学習していく。そのために、各重みパラメータに関する損失関数の勾配を求め、重みパラメータを勾配方向に微少量だけ更新する。この時の微少量を学習率と呼ぶ。

3.1.3 パラメータの最適化

3.1.2節での重みパラメータ最適化方法は確率的勾配降下法（SGD）と呼ばれ、以下のように数式で表せる。

$$\mathbf{W} \leftarrow \mathbf{W} - \eta \frac{\partial L}{\partial \mathbf{W}}$$

ここで η は学習率を表す。SGD は単純で実装も簡単だが、関数の形状が等方的でないと非効率な経路で探索することになる。そのため、問題によっては SGD よりも効率的な最適化法が存在する。ここでは最適化法の例をいくつか紹介する。

Momentum

Momentum という手法は数式で以下のように表される。

$$\begin{aligned} \mathbf{v} &\leftarrow \alpha \mathbf{v} - \eta \frac{\partial L}{\partial \mathbf{W}} \\ \mathbf{W} &\leftarrow \mathbf{W} + \mathbf{v} \end{aligned}$$

SGD と比べると更新経路のジグザク度合いが軽減される。

AdaGrad

AdaGrad はパラメータの要素ごとに適応的に学習率を調整しながら学習を行う手法。AdaGrad の更新方法は数式で以下のように表される。

$$\begin{aligned} \mathbf{h} &\leftarrow \mathbf{h} + \frac{\partial L}{\partial \mathbf{W}} \odot \frac{\partial L}{\partial \mathbf{W}} \\ \mathbf{W} &\leftarrow \mathbf{W} - \eta \frac{1}{\sqrt{\mathbf{h}}} \frac{\partial L}{\partial \mathbf{W}} \end{aligned}$$

パラメータの要素の中で大きく更新された要素は、学習係数が小さくなるようにしている。

Adam

Momentum と AdaGrad を融合したような手法。本研究では optimizer に Adam を使用している。Adam の更新方法を数式で表すと、減衰パラメータ $\rho, \rho_f \in (0, 1)$ と t 番目の反復で学習率 α_t を用いて以下のようになる。Adam のオリジナル論文では ρ と ρ_f の初期値をそれぞれ 0.999 と 0.9 に設定することをすすめている。

$$\begin{aligned}\mathbf{A} &\leftarrow \rho \mathbf{A} + (1 - \rho) \left(\frac{\partial L}{\partial \mathbf{W}} \right)^2 \\ \mathbf{F} &\leftarrow \rho_f \mathbf{F} + (1 - \rho_f) \left(\frac{\partial L}{\partial \mathbf{W}} \right) \\ \mathbf{W} &\leftarrow \mathbf{W} - \frac{\alpha_t}{\sqrt{\mathbf{A}}} \mathbf{F} \\ \alpha_t &= \alpha \left(\frac{\sqrt{1 - \rho^t}}{1 - \rho_f^t} \right)\end{aligned}$$

3.1.4 正則化

機械学習の問題では過学習が問題になることがよくある。過学習とは訓練データだけに適応しすぎてしまい、訓練データに含まれない他のデータにはうまく対応できない状態をいう。機械学習では汎化性能が求められるため、過学習を抑制するテクニックが重要になる。

Weight decay

過学習抑制のために使われる手法に Weight decay という手法がある。Weight decay は、すべての重みに対して損失関数に $\frac{1}{2} \lambda \mathbf{W}^2$ を加算する。これは、学習の過程において大きな重みを持つことに対してペナルティを課すことで、過学習を抑制しようというもの。

Dropout

Dropout は、ニューロンをランダムに消去しながら学習する手法。訓練時に隠れ層のニューロンをランダムに選び出し、選出したニューロンを消去する。これは毎回異なるモデルを学習させていると解釈でき、アンサンブル学習と同様の効果を擬似的に一つのネットワークで実現していると考えられる。

3.1.5 ハイパーパラメータの最適化

ニューラルネットワークでは重みやバイアスといったパラメータとは別に、ハイパーパラメータが数多く登場する。ここでいうハイパーパラメータとは、例えば、各層のニューロンのサイズやバッチサイズ、パラメータ更新の際の学習率や weight decay など。そのようなハイパーパラメータは、適切な値に設定しなければ、性能の悪いモデルになってしまう。そのため、ハイパーパラメータの最適化を行うことは重要になる。

3.2 畳み込みニューラルネットワーク

ニューラルネットワークの 1 種である畳み込みニューラルネットワーク (CNN) は主に画像認識や音声認識などで利用されている。CNN は図 3.1 のような全結合層に加えて、特徴的な 2 つの層、畳み込み層とプーリ

ング層から構成される。

3.2.1 畳み込み層

全結合層では3次元データを1次元にして入力する必要がある関係上、データの形状が無視されてしまう。一方、畳み込み層は入力を3次元のデータとして受け取り、3次元のデータとして出力するため、データの形状を維持できる。

畳み込み演算

畳み込み層で行う処理は畳み込み演算と呼ばれ、画像処理でいうところのフィルター演算に相当する。2次元データに対しては2次元のフィルターを用いて積和演算をすることで出力を計算する。CNNの場合、フィル

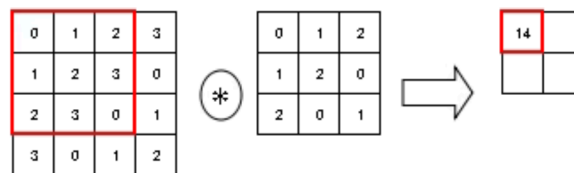


図3.3 畳み込み演算の例

タのパラメータがこれまでの重みに対応し、これを学習により最適化していく。

畳み込み演算の特徴の一つは並進等価性である。これは、畳み込み全体に渡ってフィルタのパラメータを共有しているためである。

パディング

畳み込み演算において、出力サイズが入力サイズから縮小されてしまうのを防ぐために、入力データの周囲に固定のデータ（例えば0など）を埋めるパディングという処理を行うことがある。例えば、図3.3でパディングの幅を1に設定すれば、出力サイズも(4,4)になってサイズが保たれる。

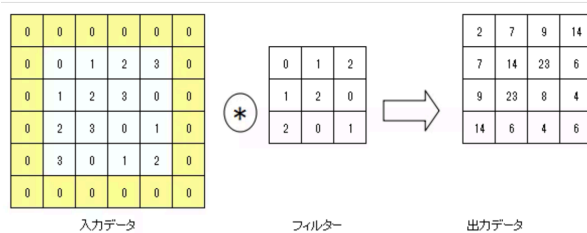


図3.4 パディングの例

ストライド

フィルターを適用する位置の間隔をストライドという。ストライドを大きくすれば、出力サイズが小さくなる。

3.2.2 プーリング層

プーリングとは、縦・横方向の空間を小さくする演算。図3.5は 2×2 の Max プーリングをストライド2で行った場合である。プーリングには Max プーリングの他に Average プーリングなどがあるが、画像認識では Max プーリングが主に用いられる。

プーリングは入力データの小さなずれに対して、同じような結果を返すため、入力データの微小なずれに対してロバストになるという特徴がある。

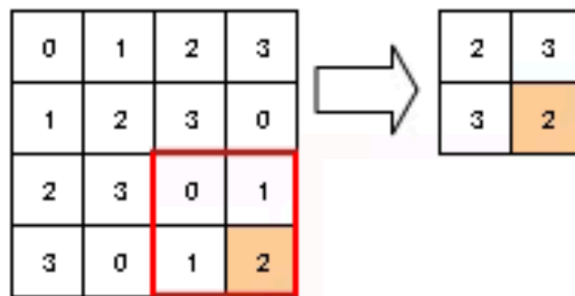


図3.5 Max プーリングの例

3.2.3 事前学習

これまで CNN の構成について話してきたが、実際の学習においてはゼロから CNN を構築することはほとんどなく、ImageNet などの膨大なデータセットですでに学習済みの CNN モデルから学習を始めることが多い。この手法によって、事前学習の段階でエッジやテクスチャなどの画像分野に汎用性の高い特徴量はすでに学習されており、自分のタスク固有の特徴量を効率的に学習することができるため、少ない訓練データかつ短時間で高精度なモデルを構築することができる。

3.3 MobileNetV3

MobileNetV3 は 2019 年に発表された CNN のモデルであり、様々なテクニックによりパラメータ数を削減することにより、軽量かつ高性能な CNN が実現された。MobileNetV3 の具体的な構成は図3.6に示している。

3.3.1 ボトルネック

ボトルネックとは残差学習で計算量を減らすためのブロック設計のこと。基本構造は

1. 1×1 畳込み チャンネル圧縮
2. 3×3 畳込み 特徴抽出
3. 1×1 畳込み チャンネル復元
4. スキップ接続 入力を出力に足す残差学習

からなる (図3.7)。このボトルネック構造により、勾配消失と効率の良い計算が達成される。

Input	Operator	exp size	#out	SE	NL	s
$224^2 \times 3$	conv2d	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	-	RE	1
$112^2 \times 16$	bneck, 3x3	64	24	-	RE	2
$56^2 \times 24$	bneck, 3x3	72	24	-	RE	1
$56^2 \times 24$	bneck, 5x5	72	40	✓	RE	2
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 3x3	240	80	-	HS	2
$14^2 \times 80$	bneck, 3x3	200	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	480	112	✓	HS	1
$14^2 \times 112$	bneck, 3x3	672	112	✓	HS	1
$14^2 \times 112$	bneck, 5x5	672	160	✓	HS	2
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	conv2d, 1x1	-	960	-	HS	1
$7^2 \times 960$	pool, 7x7	-	-	-	-	1
$1^2 \times 960$	conv2d 1x1, NBN	-	1280	-	HS	1
$1^2 \times 1280$	conv2d 1x1, NBN	-	k	-	-	1

図3.6 MobileNetV3 の仕様

Operator のBneck はBottle neck の略。SE はSqueeze and Excitation モジュールをそのブロックで用いているか否かを表す。NL は活性化関数を表し、HS はh-swish 関数、RE はReLU 関数を意味する。

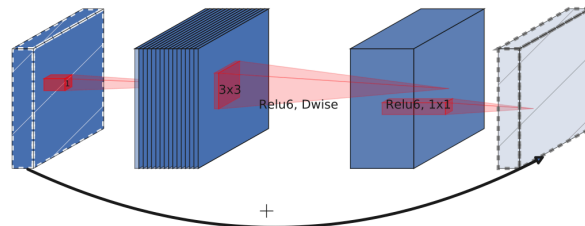


図3.7 ボトルネック構造 [10]

3.3.2 Squeeze and Excitation モジュール

以下の3つのステップからなる。

1. squeeze 各チャンネルの空間情報をグローバル平均プーリングによって1つの数に圧縮して、1次元ベクトルにする。
2. excitation 1次元ベクトルを2層の全結合層に通すことで、各チャンネルの重要度を学習する。
3. scaling 得られた重みを元の入力の各チャンネルに掛け合わせる。

このモジュールによって、入力に応じてどのチャンネルが重要であるかネットワーク自体が適応的に判断するため、モデルの精度が向上する。

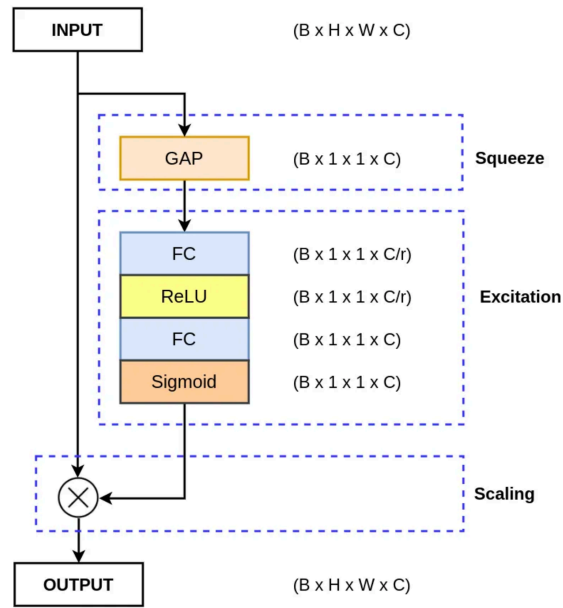


図3.8 squeeze and excitation モジュール

3.3.3 h-swish 関数

h-swish 関数とは以下で定義される関数。

$$\text{h-swish}(x) = x \frac{\text{ReLU6}(x + 3)}{6}$$

$$\text{ReLU6}(x) = \min(\max(x, 0), 6)$$

第 4 章

CNN による粒子識別の方法

4.1 物理背景

SK では大気ニュートリノや T2K 実験において、 ν_e, ν_μ を観測するため、 $\nu_e + N \rightarrow e + N$ による e イベント、 $\nu_\mu + N \rightarrow \mu + N$ による μ イベントを識別できる必要がある。また、T2K 実験によるニュートリノビームによる中性カレント相互作用 $\nu + N \rightarrow \nu + N + \pi^0$ によって生じる π^0 が崩壊して生まれる 2γ によるイベントもあるため、 π^0 イベントとの識別も必要になる。そのため、本研究では MOBILENetV3 を用いて $e/\mu, e/\pi^0, e/\gamma$ の識別を行った。本章ではシミュレーションデータの生成方法、データの preprocessing 方法、CNN による学習方法について説明を行う。

4.2 光量調整

粒子識別に先立って、粒子種以外の情報が等価になるように、異なる粒子間においてスーパーカミオカンデで検出される平均総光量が等しくなるような粒子の運動量のペアを求めたい。

4.2.1 光量調整の方法

まず、各粒子について運動量と平均総光量の関係を求める
 e と π_0 については
運動量 $p = \{100 \text{ MeV}, 200 \text{ MeV}, 300 \text{ MeV}, 400 \text{ MeV}, 500 \text{ MeV}\}$,
 μ については $p = \{300 \text{ MeV}, 400 \text{ MeV}, 500 \text{ MeV}, 600 \text{ MeV}, 700 \text{ MeV}\}$ の 5 点について平均総光量を求めて、運動量と平均総光量の関係をプロットし、1 次または 2 次関数で近似する。
そして、上で求めた関係式から各粒子ペアについて、平均総光量が一致するような運動量の関係式を求める。
最後に電子の運動量 p_e が $p_e = \{200 \text{ MeV}, 500 \text{ MeV}, 1000 \text{ MeV}\}$ の時に平均総光量が一致する μ と π_0 の運動量を求める。

4.2.2 光量調整の結果

プロットから求められた運動量と平均総光量 (potot) の関係式は

$$\begin{aligned} potot &= 9.61 (p_e/\text{MeV}) + 74.1 \\ potot &= 0.00242 (p_\mu/\text{MeV})^2 + 7.37 (p_\mu/\text{MeV}) - 1300 \end{aligned}$$

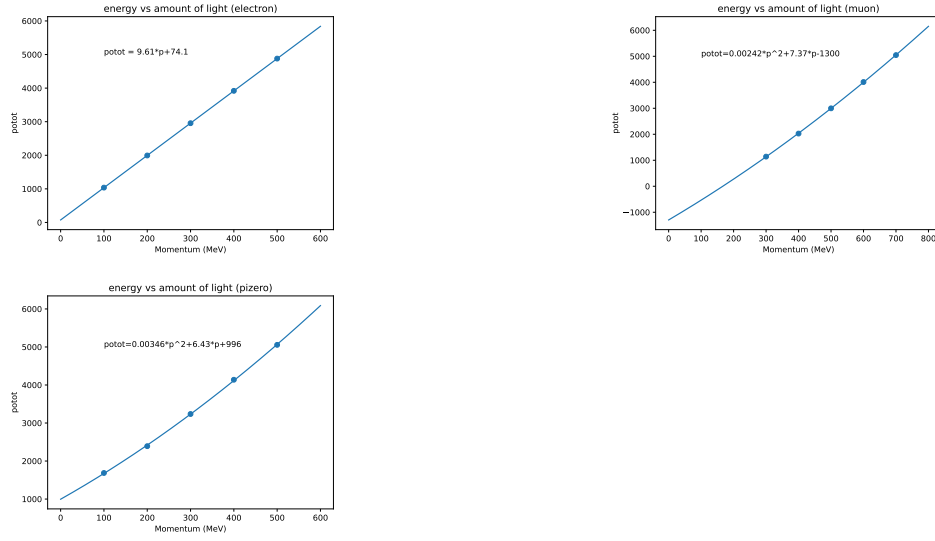


図4.1 各粒子の運動量と総光量の関係 左上： e 、右上： μ 、左下 π^0

$$potot = 0.00346 (p_{\pi_0}/\text{MeV})^2 + 6.43 * (p_{\pi_0}/\text{MeV}) + 996$$

上式から平均総光量が一致するような運動量の関係式は

e と μ について

$$(p_e/\text{MeV}) = \frac{1}{9.61} \{0.00242(p_\mu/\text{MeV})^2 + 7.37(p_\mu/\text{MeV}) - 1370\}$$

e と π_0 について

$$(p_e/\text{MeV}) = \frac{1}{9.61} \{0.00344(p_{\pi_0}/\text{MeV})^2 + 6.43(p_{\pi_0}/\text{MeV}) + 922\}$$

よって $p_e = \{200 \text{ MeV}, 500 \text{ MeV}, 1000 \text{ MeV}\}$ の時に平均総光量が一致する μ と π_0 の運動量は、

表4.1 各粒子種の平均総光量が等しくなるような運動量

粒子種 平均総光量	e	μ	π_0
2000 p.e.	200 MeV/c	396 MeV/c	146 MeV/c
4880 p.e.	500 MeV/c	685 MeV/c	480 MeV/c
9700 p.e.	1000 MeV/c	1154 MeV/c	991 MeV/c

4.3 シミュレーションデータ生成

シミュレーションデータの生成は、SKdetsim(SKのシミュレーションプログラム)で e, μ, γ, π_0 それぞれを10,000 events ずつ生成した。この時、運動量は表4.2の通り各粒子3パターンずつで、バーテックスは有効体積内(壁から2m以上離れた領域)でランダム、粒子の方向もランダムとした。

粒子種 平均総光量	e	μ	γ	π_0
2000 p.e.	200 MeV/c	396 MeV/c	200 MeV/c	146 MeV/c
4880 p.e.	500 MeV/c	685 MeV/c	500 MeV/c	480 MeV/c
9700 p.e.	1000 MeV/c	1154 MeV/c	1000 MeV/c	991 MeV/c

表4.2 生成した各粒子の運動量

ここで運動量を決定した方法は、節4.2で見たように各粒子イベントでの平均総光量が一致するようにして決定した。図4.3がSKdetsimで生成したイベントディスプレイの例である。 e は電磁シャワーを発生させるため、 μ と比べてリングの輪郭がぼやけている。 γ も最初以外は e と同じ物理過程でシャワーを発生させるため、 e と同様のリングになる。 π^0 については検出器内で2つの γ に崩壊するため、リングが2つ検出される。

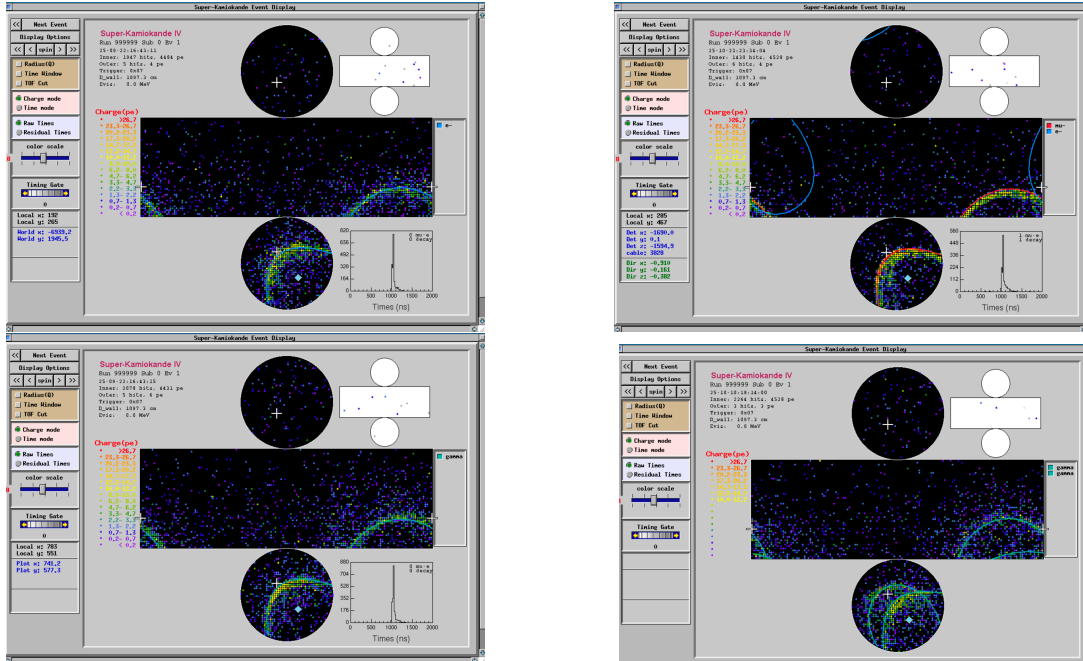


図4.2 SKdetsimで生成したイベントの例。左上が e イベント、右上が μ イベント、左下が γ イベント、右下が π_0 イベントに対応する。4つの粒子ともバーテックスは中心、粒子の運動量方向は斜め下にした

4.4 データの前処理

SKdetsimの出力ファイル（dat形式）をCSV形式に変換した後、CSV形式からhdf5形式に変換した。hdf5形式に変換する際にhit timeの情報は落として、hit chargeの情報のみをファイルに保存して、粒子識別にはhit chargeの情報のみを使用した。この際、SKの円筒展開図から 224×224 の正方形の画像に変形してCNNの入力にするために、側面はそのままに上面と底面を極座標表示によって四角に変形して側面の上下に貼り付けるという変換を行った。（図4.3、図4.4）

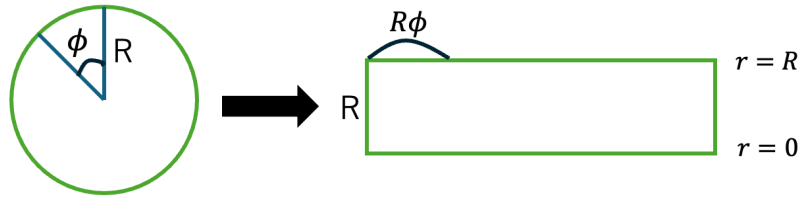


図4.3 上底面を長方形に変換する方法

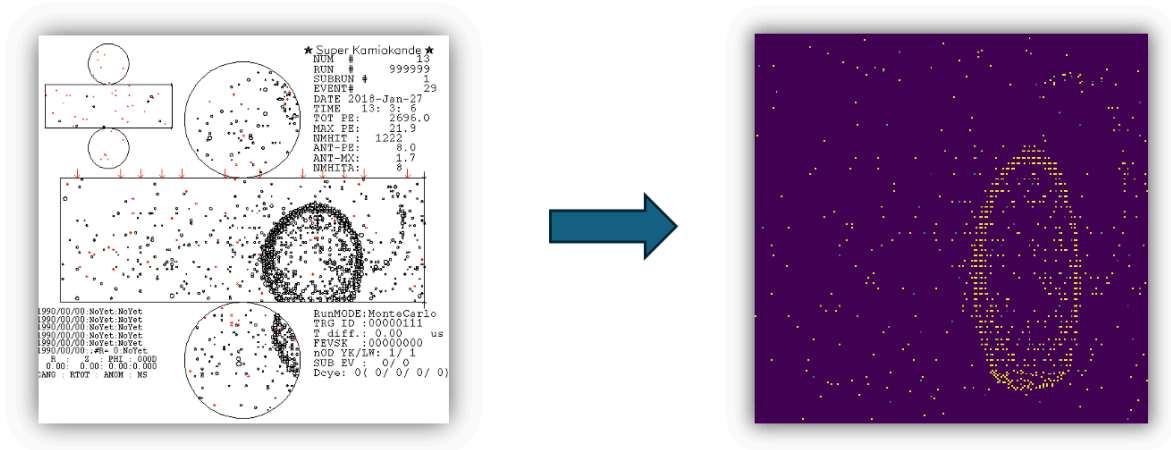


図4.4 円筒画像から正方形画像への変換

その後、それぞれ 10,000 events の異なる 2 粒子のイベントセットを 7 : 3 の割合で train data と validation data にランダムに割り振り、各粒子の train data 同士、validation data 同士を結合させた。

4.5 ハイパーパラメータの調整

CNN のモデルを決定する際には、3 章で説明したように悪いモデルにならないようにハイパーパラメータの調整を行う必要がある。ここではいくつかの異なる learning rate と weight decay の値で e と μ の識別を行なって各モデルでの accuracy を計算し、最適なハイパーパラメータを求めた。この時、 $p_e = 500$ MeV とした。また、learning rate と weight decay の値は表4.3のように変化させた。

Lr \ Wd	0.0001	0.001	0.01
0.00001	model 1	model 2	model 3
0.0001	model 4	model 5	model 6
0.001	model 7	model 8	model 9

表4.3 ハイパーパラメータの調整のための各モデル Lr は learning rate Wd は weight decay の略

得られた結果は図4.5のようであった。ここで各モデルの accuracy は 96 ~ 100 epoch の accuracy を平均した値をとった。

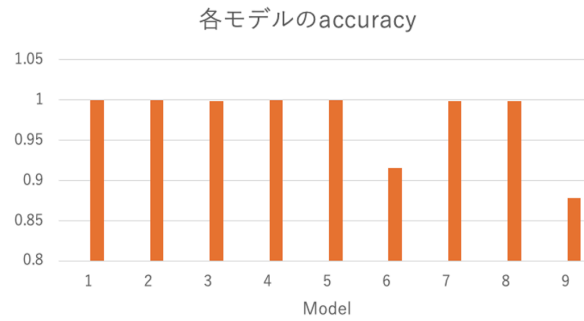


図4.5 ハイパーパラメータ調整 各モデルの accuracy

図4.5からわかるように明らかに model 6 と model 9 の性能が悪く、他のモデルはいずれも accuracy が 0.99 を超えている。このため、model 6 と model 9 に共通する learning rate の値 0.01 が悪さをしていると判断して、learning rate には一つ小さな値 0.001 を採用し、Weight decay は調べた範囲では accuracy に影響がなかったため、真ん中の値 0.0001 を採用した。すなわち、本研究ではモデル 5 のハイパーパラメータである、 $Lr = 0.001$, $Wd = 0.0001$ を選択した。

4.6 CNN による学習

表4.4に示したモデルとハイパーパラメータで train data について学習を行った。学習を行ったモデルと各ハイパーパラメータは表4.4に示されている。

model	mobilenetv3_large_100.ra_in1k
criterion	BCEwithLogitsLoss
optimizer	AdamW
batch size	32
learning rate	0.001
weight decay	0.0001
epochs	100

表4.4 CNN のモデルとハイパーパラメータ

学習の結果得られた 0 ～ 100 までの各 epoch のモデルを用いて、train data と validation data で e と μ の識別を行い、それぞれの dataset についての accuracy を計算した。ここでの accuracy の定義は

$$\frac{(e \text{ と判定された真の } e \text{ イベント数}) + (\mu \text{ と判定された真の } \mu \text{ イベント数})}{\text{全イベント数}}$$

である。また前節で見たように、learning rate と weight decay の値は調整を行なって決定した。

第 5 章

CNN による粒子識別の結果

5.1 e と μ の識別

CNN による e と μ の識別の accuracy の各 epoch における値は図5.1~5.3のようになった。また、最終的な accuracy の値は表5.1に示した。図5.1からわかるように、 $p_e = 200$ MeV/c では valid data の accuracy について、0 epoch から 0.9 とかなり高いもののおおよそ 40 epoch までは epoch ごとの accuracy のふらつきが大きく、40 epoch 以降はほとんど安定して accuracy の値は 0.98 程度で落ち着いている。一方で、 $p_e = 500, 1000$ MeV/c では、train, valid data とともに 0 epoch から accuracy が 1.0 と非常に高く 100 epoch まで安定している。運動量に関わらず accuracy に時々見られる下向きのスパイクは MobileNet 固有のものである。

$p_e(\text{MeV/c})$	$p_\mu(\text{MeV/c})$	accuracy at 100 epoch
200	396	0.98 ± 0.01
500	685	1.00 ± 0.01
1000	1154	1.00 ± 0.01

表5.1 e と μ の運動量のペアと 100 epoch でのモデルによる e/μ の識別正答率

5.2 e と γ の識別

CNN による e と γ の識別の結果は図5.4~5.6のようになった。また、100 epoch における accuracy は表5.2に示した。

e の運動量	γ の運動量	accuracy at 100 epoch
200 MeV/c	200 MeV/c	0.52 ± 0.01
500 MeV/c	500 MeV/c	0.51 ± 0.01
1000 MeV/c	1000 MeV/c	0.49 ± 0.01

表5.2 e と γ の運動量のペアと 100 epoch でのモデルによる e/γ の識別正答率

図5.4~5.6からわかるように、いずれの運動量においても train data は 0 epoch では accuracy が 0.5 であ

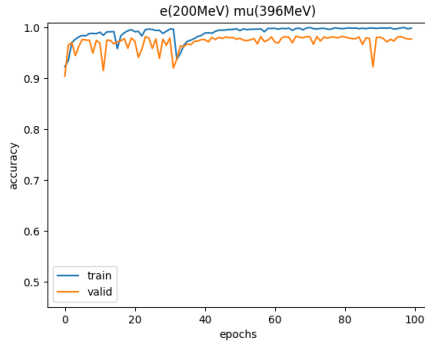


図5.1 運動量 200 MeV/c の e と運動量 396 MeV/c の μ の識別における accuracy

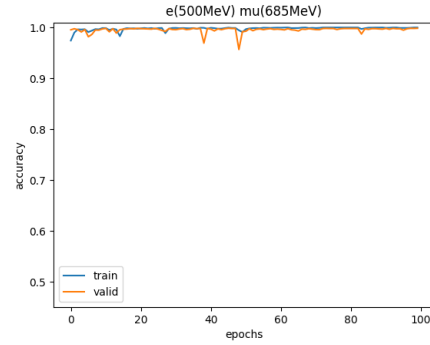


図5.2 運動量 500 MeV/c の e と運動量 685 MeV/c の μ の識別における accuracy

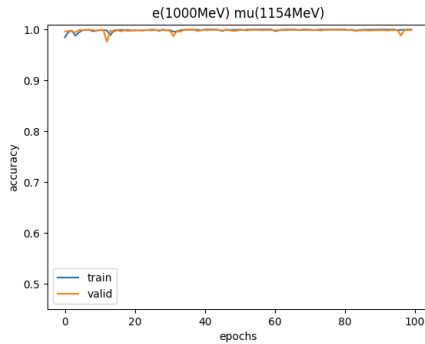


図5.3 運動量 1000 MeV/c の e と運動量 1154 MeV/c の μ の識別における accuracy

るものの 20 epoch 学習した段階では 0.9 を超え、100 epoch ではほぼ 1 となっているが、一方で valid data は最初から最後までほぼ 0.5 のままであり、学習がうまくいかず過学習が起きていると言える。

5.3 e と π^0 の識別

CNN による e と π^0 の識別の結果は図5.7~5.9のようになった。また、100 epoch における accuracy は表5.3に示した。図5.7~5.9からわかるように、train data はいずれの運動量でも accuracy の値について、0 epoch における 0.6 ~ 0.8 から学習が進み 100 epoch ではほぼ 1 を達成している一方で、valid data の accuracy は運動量 200 MeV/c では 0.95, 運動量 500 MeV/c では 0.75, 運動量 1000 MeV/c では 0.7 と train data よりも低く、 e/π^0 でも過学習が起きていると言える。また、運動量が大きくなるほど valid data の accuracy が下がっている。これは、 π^0 が崩壊するときに 2 つの γ が出て SK では 2 リングが観測されるが、 π^0 の運動量が高いほど 2 つの γ の運動量方向が一致しやすく、すなわち 2 つのリングが重なりやすくなるために、 e と一つの γ の識別に近づき、識別の accuracy が低くなると考えられる。参考のために、 $p_e = 500$ MeV/c の時の e, π^0 それぞれの誤識別率は表5.4のようであった。ただし、この結果は epoch によって少し安定しないため 96~100 epoch の平均値をとった。

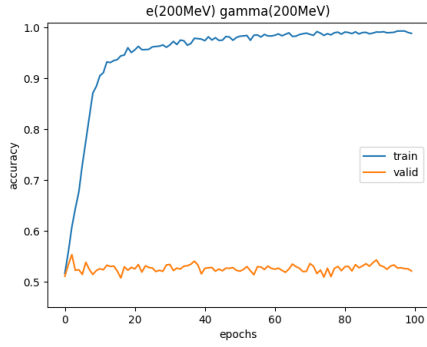


図5.4 運動量 200 MeV/c の e と運動量 200 MeV/c の γ の識別における accuracy

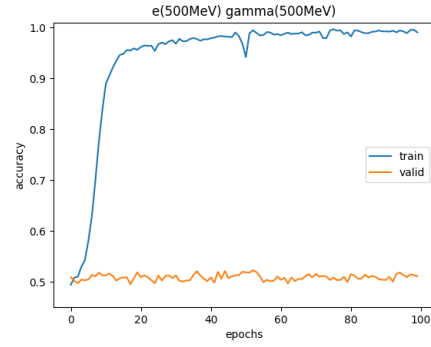


図5.5 運動量 500 MeV/c の e と運動量 500 MeV/c の γ の識別における accuracy

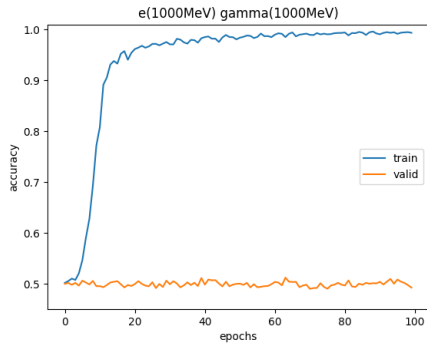


図5.6 運動量 1000 MeV/c の e と運動量 1000 MeV/c の γ の識別における accuracy

e の運動量	π^0 の運動量	accuracy at 100 epoch
200 MeV/c	146 MeV/c	0.95 ± 0.01
500 MeV/c	480 MeV/c	0.77 ± 0.01
1000 MeV/c	991 MeV/c	0.70 ± 0.01

表5.3 e と π^0 の運動量のペアと 100 epoch でのモデルによる e/π^0 の識別正答率

e	π^0
0.74 ± 0.01	0.81 ± 0.01

表5.4 e, π^0 それぞれの accuracy $p_e = 500$ MeV/c での 96~100 epoch での平均値

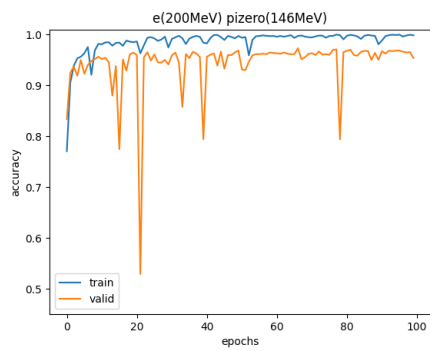


図5.7 運動量 200 MeV/c の e と運動量 146 MeV/c の π^0 の識別における accuracy

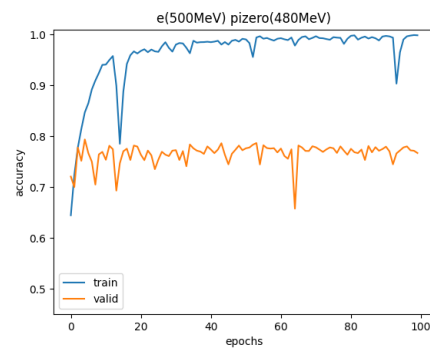


図5.8 運動量 500 MeV/c の e と運動量 480 MeV/c の π^0 の識別における accuracy

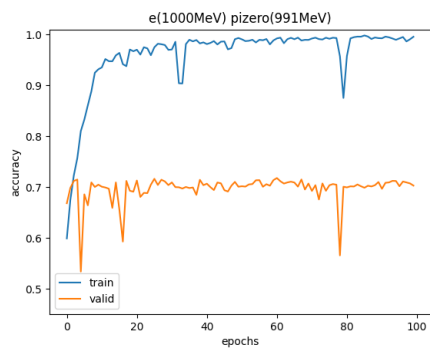


図5.9 運動量 1000 MeV/c の e と運動量 991 MeV/c の π^0 の識別における accuracy

第 6 章

考察

6.1 fiTQun と CNN の性能比較

6.1.1 accuracy の比較

CNN による粒子識別の精度を fiTQun による粒子識別の精度と比較すると表6.1のようになる。表6.1か

$p_e(\text{MeV}/c)$	e/μ by CNN	e/μ by fiTQun	e/π^0 by CNN	e/π^0 by fiTQun	e/γ by CNN
200	0.98 ± 0.01	0.94 ± 0.01	0.95 ± 0.01	0.96 ± 0.01	0.52 ± 0.01
500	1.00 ± 0.01	0.98 ± 0.01	0.77 ± 0.01	0.91 ± 0.01	0.51 ± 0.01
1000	1.00 ± 0.01	1.00 ± 0.01	0.70 ± 0.01	0.77 ± 0.01	0.49 ± 0.01

表6.1 CNN と fiTQun による粒子識別の accuracy の比較

らわかるように e/μ の accuracy は CNN の方が良い一方で、 e/π^0 の accuracy は $p_e = 500, 1000 \text{ MeV}/c$ で fiTQun の方が 0.1 程度高くなっている。また、CNN による e/γ の識別はうまくいかなかった。CNN と fiTQun での accuracy の違いは、おそらく画像認識の精度は CNN の方が高いために PMT の電荷の情報だけで識別できる場合は CNN の方が accuracy が高くなる一方で、今回 CNN には含んでいない時間の情報を fiTQun が含んでいるために、電荷の情報だけでは区別しにくく時間の情報が識別に必要な場合は fiTQun の方が accuracy が高くなると考えられる。すなわち、 e/μ は電荷だけでほとんど識別できるため CNN の方が accuracy が高く、 e/π^0 の γ の 2 リングがほとんど重なっているイベント群は電荷だけでは識別が難しいため fiTQun の方が accuracy が高くなったと考えられる。

6.1.2 CNN による識別に要する時間

ここでは、MobileNetV3 を用いた $e/\mu, e/\pi^0$ 識別の評価にかかる時間を $p_e = 200, 500, 1000 \text{ MeV}/c$ について測定した。この時 GPU は Nvidia A100 40GB を用いた。測定結果は表6.2のようであった。また、各モデルの学習にかかった時間は約 10 時間程度 (100 epoch) であった。

6.1.3 計算時間の比較

識別にかかる時間についても CNN と fiTQun で比較すると表6.3のようである。fiTQun については粒子識別に再構成も含まれているため単純な時間の比較は難しいが、粒子を識別するという用途に限れば CNN の方

$p_e(\text{MeV}/c)$	e/μ の識別時間 (s/events)	e/π^0 の識別時間 (s/events)
200	0.020	0.020
500	0.024	0.020
1000	0.020	0.011

表6.2 CNN による計算時間

p_e (MeV/c)	e/μ by CNN (s/event)	e/π^0 by CNN (s/event)	fiTQun (s/event)
200	0.020	0.020	85.6
500	0.024	0.020	136.8
1000	0.020	0.011	177.9

表6.3 CNN と fiTQun による粒子識別に要する時間の比較

が圧倒的に計算時間が短いことがわかる。

6.2 光量の角度分布

6.2.1 粒子による光量の角度分布の違い

本節では粒子ごとの光量の角度分布の違いから、粒子識別での accuracy の違いの説明を試みる。ここでの角度とは粒子の運動量方向とバーテックスから各 PMT を結んだ直線がなす角のことを意味する。また、各 PMT で検出された電荷を光量とみなしている。図6.1が光量の角度分布について電子と各粒子での比較をした図である。ここで青線が電子、赤線が他の粒子に対応している。右上の e と μ の図では、 μ の方が 42° 付近の光量のエッジがはっきりしており、 42° 以上の角度で e と μ の光量の違いがはっきりとわかる。一方で、左下の e と γ の図では互いの光量分布がほとんど重なっており違いが見られない。また、右下の e と π^0 の図については、 e の方がピークが鋭くなっているものの、ピーク以外の部分では光量に大きな違いはない。

以上より、 e と μ では広い角度範囲で光量に違いがあるため CNN によってほぼ完全に識別することができたが、 e と γ では光量の角度分布全体が一致しているために CNN によっても識別することができなかったと説明できる。また、 e と π^0 は角度光量分布のピーク位置での光量の差はあるが、 e と μ ほど広い角度範囲の差ではないために CNN による粒子識別の accuracy が 75 % 程度に収まってしまったと考えられる。

6.2.2 運動量による光量の角度分布の違い

本節では e と π^0 の光量角度分布について、運動量 $(p_e, p_{\pi^0}) = \{(200 \text{ MeV}/c, 146 \text{ MeV}/c), (500 \text{ MeV}/c, 480 \text{ MeV}/c), (1000 \text{ MeV}/c, 991 \text{ MeV}/c)\}$ の3組での違いから粒子識別の accuracy の違いの説明を試みる。図6.2が角運動量での光量角度分布を表している。図からわかるように運動量が高いほど互いの光量角度分布の差が小さくなっており、これは CNN による粒子識別の accuracy が運動量が高くなるほど低くなることに対応しているといえる。

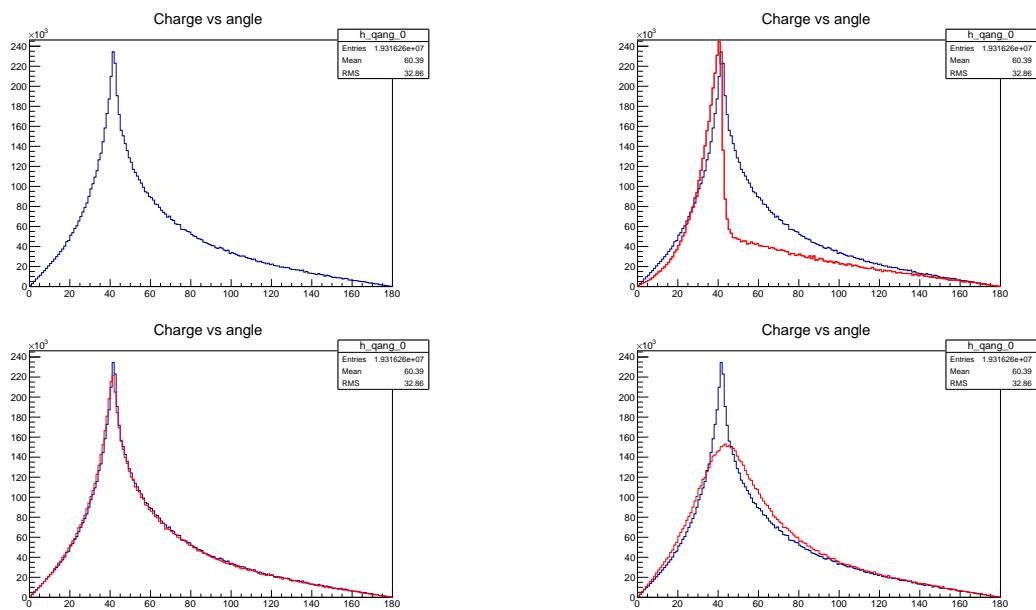


図6.1 e の光量角度分布（青線）と μ, γ, π^0 の光量角度分布（赤線） 左上： e のみ、右上： e と μ 、左下： e と γ 、右下： e と π^0 運動量は $(p_e, p_\mu, p_\gamma, p_{\pi^0}) = (500 \text{ MeV/c } 685 \text{ MeV/c } 500 \text{ MeV/c } 480 \text{ MeV/c})$

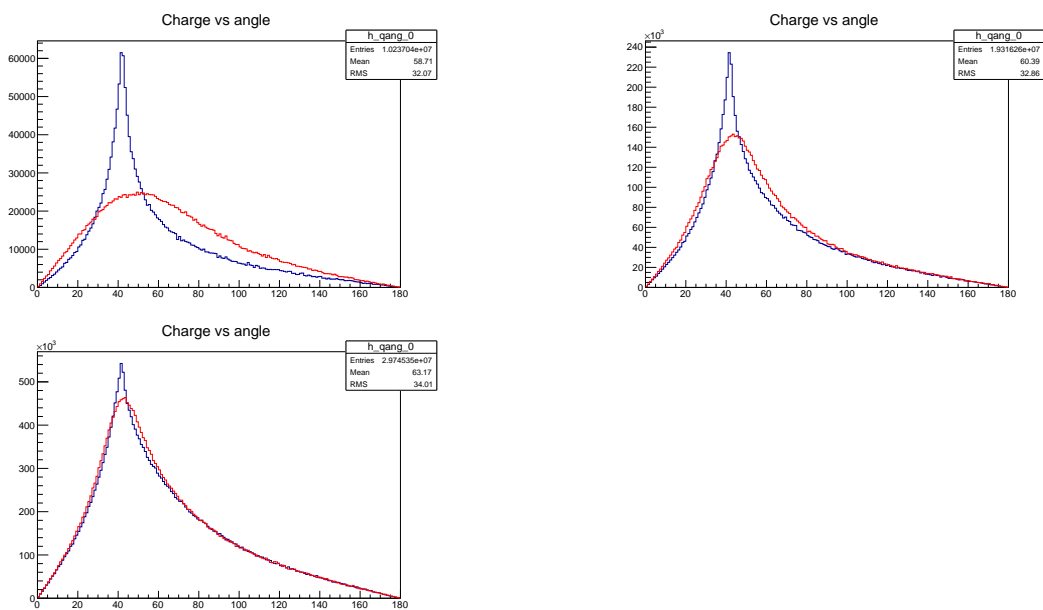


図6.2 e の光量角度分布（青線）と π^0 の光量角度分布（赤線） 左上： $p_e = 200 \text{ MeV/c}$ 右上： $p_e = 500 \text{ MeV/c}$ 左下： $p_e = 1000 \text{ MeV/c}$

第 7 章

まとめ

本研究では SK での粒子識別の精度および計算時間の向上を目的として、CNN の 1 モデルである MobileNetV3 を用いて SK のシミュレーションデータにおける e/μ , e/γ , e/π^0 の粒子識別を行った。シミュレーションデータは各粒子 10000 events ずつ生成し、MobileNetV3 によって 100 epochs の学習を行なった。学習結果は e/μ の識別率が 1.0 程度まで識別でき、 e/π^0 の識別率は 70% から 90% で運動量上がるほど識別率が下がった。 e/γ については識別率が約 0.5 と識別がうまくいかなかった。また、識別にかかった時間は 1event あたり約 0.02 秒であった。

CNN による識別結果を fiTQun の結果と比較すると、accuracy については e/μ は CNN が上回ったが、 e/π^0 は CNN が下回る結果となった。識別にかかる時間は CNN の方が約 1/10000 と非常に速い結果になった。

また、本研究で光量の角度分布から粒子識別における accuracy の違いの説明を試み、実際光量の角度分布から粒子の違いおよび運動量の違いによる accuracy の差を定性的に説明をすることができた。

以上から、CNN はこれまで用いられてきた fiTQun に比べて非常に短い時間で粒子識別を行うことができると言える。そのため、より計算コストの低いプログラムが必要となるハイパーカミオカンデ実験において CNN を用いたイベントの解析を行うことが考えられる。

今後の展望としては、より計算コストが高い CNN である ResNet などのモデルを用いることで accuracy の向上を図ることが期待される。粒子識別以外に運動量やバーテックスも機械学習によって求めることで、fiTQun が行なっている事象再構成を置き換えることができることも考えられる。

参考文献

- [1] Q. R. Ahmad et al. (SNO Collaboration), "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory," *Physical Review Letters* 89, 011301 (2002)
- [2] Y. Fukuda et al.. "Evidence for Oscillation of Atmospheric Neutrinos" . *Phys. Rev. Lett.*, 81:1562, 1998.
- [3] K. Abe et al. (T2K Collaboration), "Observation of electron neutrino appearance in a muon neutrino beam," *Phys.Rev.Lett.*, vol.112,no.6,p.061812,2014.
- [4] K. Abe, S. Abe, C. Bronner, Y. Hayato, K. Hiraide, K. Hosokawa, et al., "First Joint Oscillation Analysis of Super-Kamiokande Atmospheric and T2K Accelerator Neutrino Data," *Phys. Rev. Lett.*, vol. 134, no. 1, p. 011801, Jan. 2025, doi:10.1103/PhysRevLett.134.011801.
- [5] S. Fukuda et al.. "The Super-Kamiokande detector" . *Nucl. Instrum. Meth.*, A501:418-462, 2003.
- [6] H. Nishino, K. Awai, Y. Hayato, S. Nakayama, K. Okumura, M. Shiozawa, A. Takeda, K. Ishikawa, A. Minegishi, and Y. Arai. "High-speed charge-to-time converter ASIC for the Super-Kamiokande detector," *Nucl. Instrum. Meth. A*, Vol. 610, pp. 710-717, 2009.
- [7] S. Yamada, et al. "Commissioning of the new electronics and online system for the Super-Kamiokande experiment," *IEEE Trans. Nucl. Sci.* Vol. 57, pp. 428-432, 2010.
- [8] Y. Nakano, T. Hokama, M. Matsubara, M. Miwa, M. Nakahata, T. Nakamura, H. Sekiya, Y. Takeuchi, S. Tasaka, and R. A. Wendell." Measurement of the radon concentration in purified water in the Super-Kamiokande IV detector," *Nucl. Instrum. Meth. A*, Vol. 977, p. 164297, 2020.
- [9] M. Jiang, K. Abe, C. Bronner, Y. Hayato, M. Ikeda, K. Iyogi, et al., "Atmospheric neutrino oscillation analysis with improved event reconstruction in Super-Kamiokande IV," **Prog. Theor. Exp. Phys.**, vol. 2019, no. 5, p. 053F01, May 2019
- [10] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3," in **Proc. 2019 IEEE/CVF Int. Conf. Comput. Vis. (ICCV)**, 2019, pp. 1314-1324,
- [11] S. Goto, "Evaluation of 50-cm photomultiplier tube stability and proton decay search sensitivity for the Hyper-Kamiokande experiment," Master's thesis, University of Tokyo, 2025
- [12] S. Tobayama, "An Analysis of the Oscillation of Atmospheric Neutrinos," Ph.D. thesis, University of Tokyo, 2010