修士論文

ILC衝突点ビーム形状モニターへの機械学習の応用 研究

 \sim The application of machine learning to an interaction-point beam profile monitor for the ILC \sim

東北大学大学院理学研究科 物理学専攻

小林 陽介

平成30年

国際リニアコライダー (International Linear Collider; ILC) は電子ビームと陽電子ビームが衝突する全長約 20km の線形型加速器である。衝突する粒子は共に内部構造を持たない粒子であるため、衝突時に発生する背景事象が少ない。ゆえにヒッグス粒子の精密測定などが期待されている。

線形型加速器は円形加速器と違いビームの衝突は1度だけである。そして、衝突頻度を上げるためビームサイズは極限まで絞られる。つまり、ILCにはビームサイズがナノメートルオーダーとなる中で正確に衝突させることが求められる。この条件を達成するためにビームに関するパラメータを測定し、加速器へフィードバックさせる必要がある。ビームパラメータを測定するためにペアモニター考案された。ペアモニターはビーム衝突時に発生する低エネルギーの電子・陽電子を利用する。そのため、ビームを非破壊で測定可能である。加えて、大量の電子・陽電子が発生するため、短時間で十分な統計量をためることができ、即時フィードバックすることも可能である。

本研究はビームパラメータ測定のため、ペアモニターの解析に機械学習を用いた新たな手法を提案することを目的としてシミュレーションによる検討を行った。本論文はペアモニターのヒット分布の解析と、機械学習を用いたビームパラメータ再構成の手法を考案する。

ビーム衝突シミュレーションで、陽電子ビームのビームサイズを変えて、ペアモニター上の電子・陽電子のヒット分布がどのように変化するか調べた。このヒット分布の変化を機械学習に通して読み取ることでビームパラメータの再構成を行う。深層学習の分野で広く用いられている Tensorflow と Keras を用いて機械学習のニューラルネットワークを作成しビームサイズの再構成を試みた。ペアモニターのデータからビームパラメータを再構成するのに機械学習が有効であることがわかった。しかし、複数のビームパラメータを正確に再構成するには、さらなる改善が必要である。

目 次

目次		ii
図目次	ζ	\mathbf{v}
表目次	र	vii
第1章	研究の目的	1
第2章	国際リニアコライダー(ILC)	2
2.1	概要	
2.2	ILC	
	2.2.1 電子源	3
	2.2.2 陽電子源	3
	2.2.3 減衰リング	3
2.0	2.2.4 主線形加速器	4
2.3	ILD	4
	2.3.1 Particle Flow Algorithm	5 5
	2.3.2 飛動機出稿	
	2.3.4 前方検出器	7
第3章	Pair Monitor	9
3.1	概要	9
3.2	測定原理	10
		12
4.1	ビームパラメータ	12
4.2	ビーム衝突シミュレーション	
	4.2.1 CAIN	
	4.2.2 Incoherent Pair のトラッキング	13
	4.2.3 シミュレーションの結果	14
第5章	解析	17
5.1	nominal なパラメータの場合の解析	17
	5.1.1 Incoherent Pair の生成	17

	1 1/4
Н	7/77
	1 1 /\

謝辞		63
第8章	まとめ	61
	7.2.3 結果	57
	7.2.2 学習ネットワーク構成	
	7.2.1 トレーニングデータ	
7.2	ビームパラメータ (σ_x, σ_y) の再構成	
- 0	7.1.5 比較	
	7.1.4 結果	
	7.1.3 畳み込みニューラルネットワークを用いた検討	
	7.1.2 3層ニューラルネットワークを用いた検討	
	7.1.1 トレーニングデータ	
7.1	ビームパラメータ (σ_y) の再構成 \dots	
第7章	機械学習への応用	45
0.5	以川 / / / 上 /	44
6.5	6.4.2 Batch Normalization	
	6.4.1 Dropout	43
6.4	精度改善の手法	43
6.4	6.3.2 プーリング層 (Pooling layer)	40
	6.3.1 畳み込み層 (Convolution layer)	39
6.3	畳み込みニューラルネットワーク	39
<i>a</i> o	6.2.5 データセット	39
	6.2.4 損失関数 (Loss function)	37
	6.2.3 活性化関数 (Activation)	36
	6.2.2 ニューラルネット	35
	6.2.1 パーセプトロン	
6.2	機械学習	34
6.1	概要	34
第6章	機械学習	34
5.4	先行研究との比較・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	30
F 4	5.3.2 nominal な場合との比較	30
	5.3.1 Incoherent Pair の解析	28
5.3	ペアモニター上ヒット分布の σ_y 依存性	28
	5.2.2 nominal な場合との比較	25
	5.2.1 Incoherent Pair の解析	25
5.2	ペアモニター上ヒット分布の σ_x 依存性 \dots	24
	5.1.3 ペアモニター上でのヒット分布	24
	5.1.2 Incoherent Pairの解析	19

目次

参考文献 64

図目次

2.1	ILC 概要図 [1]	2
2.2	陽電子源 [1]	3
2.3	9 セル超伝導加速空洞 [1]	4
2.4	ILD 概要図 [1]	5
2.5	Vertex Detector 機械的構造概要図 [1]	6
2.6	Silicon Tracker[1] \dots	7
2.7	前方検出器 [2]	8
3.1	3つの Incoherent 過程	10
3.2	電子陽電子対生成の様子	10
4.1	座標	12
4.2	鉛直方向のずれ δ_y	13
4.3	Incoherent Pair 生成の時間発展	15
4.4	Incoherent Pair 分布	16
4.5	Incoherent Pair 分布 (拡大図)	16
5.1	ペアモニターの位置	17
5.2	Incoherent Pair を見る位置の定義	18
5.3	$arphi$ 分布 \ldots	18
5.4	ペアモニター上での Incoherent Pair	19
5.5	Incoherent Pair(電子) を解析的に解いた結果 $(p_t \mathbb{N})$	20
5.6	電子がペアモニターに衝突するまでの回転数	21
5.7	Incoherent Pair(陽電子) を解析的に解いた結果 $(p_t \mathbb{N})$	22
5.8	陽電子がペアモニターに衝突するまでの回転数	23
5.9	ペアモニターヒット分布図 (nominal パラメータ)	24
5.10	ペアモニターヒット分布図 (電子陽電子別)	25
5.11	Incoherent Pair $(\sigma_x = 1.5\sigma_x^{TDR})$	26
	φ 分布の σ_x 比較 \dots	26
	ペアモニターヒット分布図 $(\sigma_x = 1.5\sigma_x^{TDR}$ 、電子陽電子別 $)$	
5.14	σ_x を変化させたときのペアモニターヒット分布図 \dots	27
	動径方向のヒット分布	
	Incoherent Pair(3.0 = σ_y/σ_y^{TDR})	29
	φ 分布の σ _n 比較	29

	g - \cdot \cdot \cdot	30
	$\sim \sim $	31
		32
		33
5.22	anti-DID による効果	33
6.1	パーセプトロン概要図 (2 入力、1 出力)	35
6.2		36
6.3		38
6.4		40
6.5		41
6.6		42
6.7		42
6.8		43
6.9		43
7.1		47
7.2		47
7.3		49
7.4		50
7.5	** * ***	50
7.6		51
7.7	ビームサイズ確率分布	52
7.8	g	53
7.9	逆行列法との比較	55
7.10		56
7.11		57
7.12	損失関数のグラフ	58
		59
7.14	ビームサイズ σ_y 再構成 \dots	60
8.1	ビームサイズ σ_y 再構成	62
8.2	複数のビームパラメータ (σ_x,σ_y) 再構成結果 \dots \dots \dots \dots	62

表目次

4.1	TDR2013 ビームパラメータ	12
5.1	ビームパラメータ	32
7.1	データセットの枚数 (3 層ニューラルネットワーク)	46
7.2	データセットの枚数	48
7.3	期待値	54

第1章 研究の目的

ILC は直線軌道上で電子ビームと陽電子ビームが衝突する線形型加速器である。線形加速器の一番の特徴は1組のビームに対して衝突の機会が一度しかないという点である。ゆえに多くの素粒子反応を観測するには単位時間あたりの反応数を上げることが重要になる。衝突する機会が一度しかなく、加えてビームサイズがナノメートルオーダーのビームを正確に衝突させることがILC に求められる。この条件を達成するためにビームサイズやビーム同士のズレといったビームに関するパラメータ測定が必要である。ペアモニターはビームパラメータを測定するための検出器として考案された。

ペアモニターにヒットした粒子は対向するビームの電磁場によって散乱されるため、ヒットした粒子の分布から相手のビームのパラメータを読み取ることができる。ヒット分布から読み取る事ができる情報にはビームの状態によって大きく値が変化する情報もあれば変化の割合が小さい情報もある。これまで行われてきた先行研究では、変化の割合が小さい情報は見つけ出すことが困難だったために大きく値が変化する情報のみを利用してビームのパラメータを取得していた。取得したいパラメータが少ない場合、例えば、ビームのビームサイズのみを測定したい場合ではこれまでの先行研究で開発された手法が有効であることが示されている。しかし、測定を行いたいビームパラメータが多くなるとその手法が有効であるとは限らない。それぞれのビームパラメータがペアモニターで読み取ることができる情報を複雑に変化させるため、人為的に選ぶのは難しい。情報の変化の割合によらず、ビームパラメータを取得する手法が求められている。

本研究は、情報の変化の割合によらずヒット分布から得られる情報をすべて用いてビーム情報を測定する方法の開発を本研究の目的とする。

本論文は、2章でILCについてILD検出器に関する概要を述べる。3章でペアモニターについて説明する。6章で本研究で用いた機械学習について述べる。4章でシミュレーション環境および使用したソフトウェアについて説明し、5章でペアモニターヒット分布の解析を行う。7章で機械学習を用いたビームパラメータの再構成手法について述べる。最後に8章でまとめを述べる。

第2章 国際リニアコライダー(ILC)

2.1 概要

国際リニアコライダー (ILC; International Linear Collider) は電子・陽電子衝突型の線形加速器 (図 2.1) である。全長 20km で重心系エネルギー 250 GeV から 500 GeV まで粒子が加速される。ILC の全長を 30 km まで拡張させ重心系エネルギーを 1 TeV まで加速させるアップグレードも予定されている。

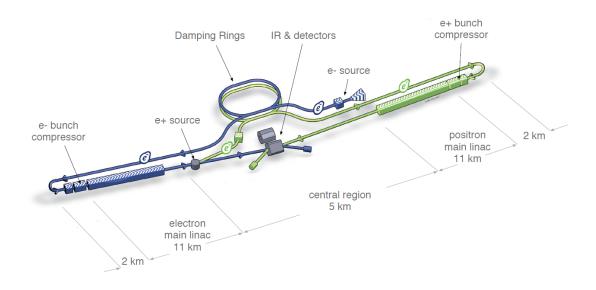


図 2.1: ILC 概要図 [1]

ILC は内部構造を持たない電子と陽電子の衝突であるため、衝突した際に発生する背景事象が陽子衝突型加速器である Large Hadron Collider (LHC) と比べて少ない。つまり、高精度で物理解析を行う事が可能になる。ILC の主な目的はヒッグス粒子やトップクォークの精密測定、新粒子探索などである。

本章では目的達成のために必要な ILC の主要設備について述べる。

2.2 ILC

ILC は主に電子・陽電子源、減衰リング、主線形加速器の3つの要素で構成される。電子・陽電子源によってビームが生成され、減衰リングでビームの広がりを抑えて収束させる。そして、主線形加速器でビームが目的のエネルギーまで加速される。

2.2.1 電子源

電子源では衝突ビームに必要な電子の生成を行う。DC 銃内部で編曲レーザが GaAs 光電陰極 (Photocathode) に照射すると、光電効果により電子が生成される。その電子は 90%の偏極が必要とされる。生成された電子は常伝導加速器によって 2.0×10^{10} 個集められバンチを形成し、 $76\,\mathrm{MeV}$ まで加速される。その後、超電導リニアックによって $5\,\mathrm{GeV}$ まで加速し、超電導ソレノイドによってスピンの向きを垂直に調整され減衰リングに送られる。

以降は多数の粒子が含まれるバンチをビームと呼ぶ。

2.2.2 陽電子源

陽電子源の主要な構成要素を図2.2に示す。

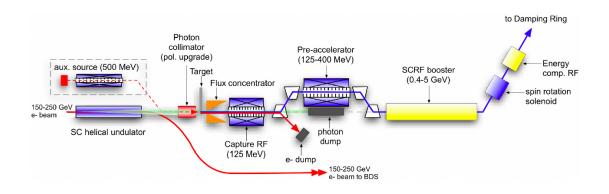


図 2.2: 陽電子源 [1]

主線形加速器で加速された電子はアンジュレータ (undulator) と呼ばれる装置の中を進む。アンジュレータは電子を蛇行させて放射光を発生させる装置である。磁場の向きが異なる磁石を交互に並べることでその中に侵入した電子は蛇行しながら運動し、制動放射によって 10 MeV から 30 MeV までのエネルギーを持つ光子が生成される。電子と光子は分離され、電子は衝突点に運ばれる。光子はチタン合金のターゲットに衝突し、大量の電子・陽電子ペアを生成する。電子・陽電子ペアは 125 MeV まで加速され、陽電子は電子と光子から切り離され、そして、陽電子は常伝導加速器によって 400 MeV まで加速される。最終的に超伝導加速器によって 5 GeV まで加速され、エネルギーとスピンの向きが揃えられて減衰リングに送られる。スピンの偏極率は 30%としている。

2.2.3 減衰リング

減衰リングの目的はビームの広がりを抑えることである。ILC は直線型であるため衝突回数が1度しかない。そのため電子・陽電子ビームは低いエミッタンスであることが要

求されている。エミッタンスとはビームの位置の広がりと運動量の広がりの積で表され、 ビームの質を決定づけるパラメータとなる。

減衰リングは周長 3.2 km で陸上競技のトラックのような形状をしている。円弧部分で電子・陽電子の各粒子は制動放射により光子が生成されてエネルギーを失う。ただし、光子は各粒子の進行方向に放射されるため、進行方向は変わらずエネルギーの絶対値のみが減少する。そして、直線部分で粒子が基準軌道方向に加速することで横方向の運動量が減少されビームの広がりが抑えられる。

2.2.4 主線形加速器

減衰リングからはじき出された電子・陽電子は Ring to Main Linac (RTML) を通過し主線形加速器に運ばれる。RTML で電子・陽電子はエネルギーの広がりを最小限に抑えながら $5 \, \text{GeV}$ から $15 \, \text{GeV}$ まで加速される。重心系エネルギーを $500 \, \text{GeV}$ とする場合、全長約 $11 \, \text{km}$ の主線形加速器によって電子・陽電子は $250 \, \text{GeV}$ まで加速される。加速には図 $2.3 \, \text{に示すニオブ製}$ $9 \, \text{セル超伝導加速空洞が約}$ $7400 \, \text{個使用され}$ 、平均加速勾配は運転時に $31.5 \, \text{MeV/m}$ となる。



図 2.3: 9 セル超伝導加速空洞 [1]

2.3 ILD

標準理論を超える新物理の探索に応えるため、ILC は Silicon Detector (SiD) と International Large Detector (ILD) という 2 つの検出器が提案 [2] されている。SiD と ILD は将来的に 1TeV までアップグレードした場合も対応できるように設計している。

本研究の対象であるペアモニターはILDで使用する検出器であるため、本節はILD内の各検出器について述べる。図 2.4 にILDの概要図を示す。ILDに含まれる主な検出器は衝突点から順に飛跡検出器、カロリメータ、ミューオン検出器があり、それぞれ役割が異なる。ILDのビーム軸に沿って、その両端にビームパラメータとルミノシティを測定する検出器が設置されており、それらの検出器をまとめて前方検出器と呼んでいる。

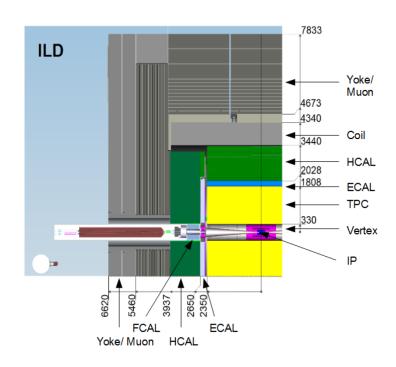


図 2.4: ILD 概要図 [1]

2.3.1 Particle Flow Algorithm

ILCで電子と陽電子が衝突したときのイベントには多数のジェットが含まれる。このジェットのエネルギーや崩壊点が精度良く見積もられるようにすることが物理解析を行う上で重要となる。Particle Flow Algorithm (PFA) [3] はジェットを構成する粒子のエネルギーや運動量を効率良く ILCで測定するために開発された測定方法である。PFA によって荷電粒子は飛跡検出器、光子は電磁カロリメータ、中性ハドロンはハドロンカロリメータでそれぞれ測定され、それらの結果を足し合わせることで高いジェットエネルギー分解能が得られる。

2.3.2 飛跡検出器

ILD の飛跡検出器は崩壊点検出器 (Vertex Detector; VTX)、シリコン飛跡検出器 (Silicon Tracker)、中央飛跡検出器 (Time Projection Chambar; TPC) から構成されている。

- 空間分解能が衝突点付近で3μm以下
- 最内層の半径が 1.6 cm
- ピクセル占有率が数%以下

VTX は図 2.5 に示すように 3 層の円筒形の構造を持つ。再構成の精度を上げるため衝突点近傍に VTX を設置する必要があるが、近いほどバックグラウンドの影響が大きくなる。その結果、バックグラウンドが VTX のピクセルセンサにヒットしピクセル占有率を上げてしまうため、要求されている性能を満たせない。その対策として、ピクセルを増やして 1 つあたりのピクセル占有率を下げる方法や、ビーム衝突の間 (200 ms) に読み出しを行うことでピクセル占有率を上げない方法が検討されている。ピクセルセンサの候補として、CMOS ピクセルセンサと Fine Pixel CCD センサが開発中である。

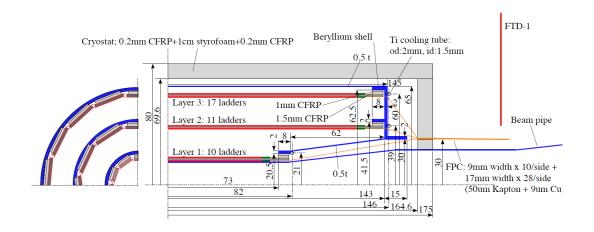


図 2.5: Vertex Detector 機械的構造概要図 [1]

シリコン飛跡検出器は4つの検出器で構成されている。

- Silicon Inner Tracker (SIT)
- Silicon External Tracker (SET)
- Endplate of the TPC (ETD)
- Forward Tracker (FTD)

4つの検出器は全て VTX の外縁部に設置されている。図 2.6 に全体図を示す。VTX と後述の TPC の間を補完して ILD 全体の運動量分解能の向上させる役割を持つ。

2.3.3 カロリメータ

ILDのカロリメータは電磁カロリメータとハドロンカロリーメータが設置されている。 どちらのカロリメータもカスケードシャワーと呼ばれる現象を利用して入射した粒子のエネルギーを測定する。電磁カロリメータはハドロンカロリメータの内側に設置され、測定する粒子は異なる。

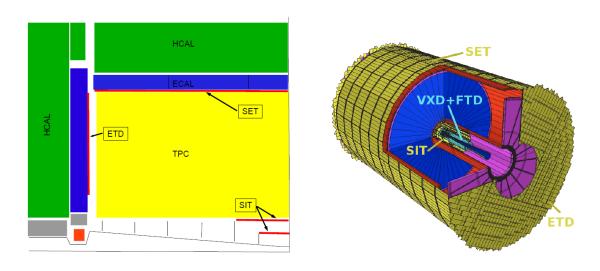


図 2.6: Silicon Tracker[1]

電磁カロリメータ

電磁カロリメータは電子、陽電子と光子のエネルギーを測定する。電磁カロリメータ内に入射した粒子はカロリメータ内の物質と相互作用し制動放射される。その結果光子が生成され、対生成によって電子・陽電子を生成する。それらの粒子は再び物質と相互作用し、最終的に大量の粒子が生成される。この一連の現象は電磁シャワーと呼ばれ、電磁カロリメータは電磁シャワーを発生させる吸収層とエネルギーを測定するセンサー層の組で構成させる。電磁カロリメータ全体の小型化と微細なエネルギー分解能が要求され、放射長 X_0 とモリエール半径 R_M が短い必要がある。モリエール半径はシャワーの横方向の広がりを表し、全シャワーの90%のエネルギーが収まる半径で定義される。吸収層にはタングステン ($X_0=3.5mm$, $R_M=9mm$) を用いた検出器が開発されている。

ハドロンカロリメータ

ハドロンカロリメータは中性ハドロンのエネルギーを測定する。ハドロンはカロリメータ内の物質の原子核と強い相互作用をするため、電磁シャワーと同じ原理によって大量の粒子が生成される。このシャワーのエネルギーは鉄で構成される吸収層とセンサー層によって測定される。

2.3.4 前方検出器

ビーム軸上に設置される複数の検出器をまとめて前方検出器と呼ばれている。主な目的はルミノシティの測定やビームサイズの測定などである。ルミノシティは LumiCal で測定され、ビームサイズは BeamCal によって測定される。各検出器の位置を図 2.7 に示す。前方検出器は約極角 $5\sim70$ mrad の範囲をカバーし、電磁カロリメータおよびハドロンカロリメータでは測定できない部分を補う。

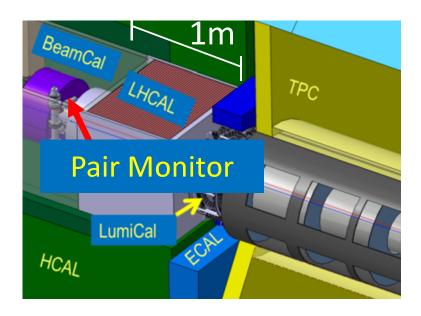


図 2.7: 前方検出器 [2]

LumiCal

LumiCal はビームが衝突した時のルミノシティを測定することを目的とする電磁カロリメータである。ルミノシティL はバーバー散乱によって測定される。バーバー散乱は電子–陽電子生成 $(e^-e^+ \to e^-e^+)$ で反応断面積 σ が大きい。LumiCal はバーバー散乱のイベント数 N_b を測定し、ルミノシティL を以下の式から見積もる。TDR 設計値では LumiCal の測定精度は 0.1% の精度が要求されている。

$$L = \frac{N_b}{\sigma} \tag{2.1}$$

BeamCal

BeamCal はビームパラメータを測定することを目的とする電磁カロリメータである。また、センサーで測定されたエネルギー損失からルミノシティを素早く見積もることも可能である。ビームパラメータおよびルミノシティは TDR 設計値で 10%の精度が要求されている。

ペアモニター

ペアモニターは衝突点のビーム形状を測定する。そして本研究で対象とする検出器である。BeamCal の前方に設置され、ILC の設計で 10%の精度でのビーム形状の再構成が要求される。次章で詳しく説明をする。

第3章 Pair Monitor

3.1 概要

ILC は前章で述べたように直線軌道上で電子ビームと陽電子ビームが衝突する線形型加速器である。線形加速器の一番の特徴は1組のビームに対して衝突の機会が一度しかないという点である。ゆえに多くの素粒子反応を観測するには単位時間あたりの反応数Nを上げることが重要になる。反応数Nは反応断面積 σ に比例する。比例定数 $\mathcal L$ と定義すると反応式は次式で表される。

$$N[s^{-1}] = \mathcal{L}[cm^{-2}s^{-1}] \times \sigma[cm^{2}]$$
(3.1)

加速器実験において比例定数 \mathcal{L} はルミノシティと呼ばれる。反応断面積は自然法則によって値が決められるが、ルミノシティは調整可能なパラメータである。つまり、反応数を上げるためにはルミノシティを大きくする必要がある。ILC 実験の成功するかどうかはルミノシティが重要な役割を持つ。

ILCで要求されるルミノシティは次式で与えられる[4]。

$$\mathcal{L} = \frac{f_{rep}n_bN^2}{4\pi\sigma_x\sigma_y}H_DH_G \tag{3.2}$$

ここで、N はバンチあたりの粒子数、 n_b はビームあたりのバンチ数、 f_{rep} は単位時間あたりのビーム衝突回数、 σ_x,σ_y が衝突点での水平、鉛直方向のビームサイズである。また、 H_D はルミノシティ増大係数と呼ばれ、 σ_x,σ_y がが電子バンチと陽電子バンチ間で互いの電磁場の影響によって実効的に収縮する度合いを表す。 H_G はビームがゼロでない交差角で衝突するとき、実際に衝突する粒子数はバンチ長 σ_z が長いほど減少する度合いを表す。ルミノシティは式 3.2 より σ_x,σ_y に反比例する。従って、要求されるルミノシティを達成するには σ_x,σ_y を小さく絞る必要がある。

衝突する機会が一度しかなく、加えてビームサイズがナノメートルオーダーのビームを 正確に衝突させることが ILC に求められる。この条件を達成するためにビームサイズや ビーム同士の位置のズレといったビームに関するパラメータ測定が必要である。ペアモ ニターはビームパラメータを測定するための検出器として考案された [5]。この検出器は ビーム衝突時のビーム・ビーム相互作用から生成される電子・陽電子を利用するため、衝 突点にモニターを設置しなくてもよい。加えて、ビーム非破壊という利点を有する。ま た、電子・陽電子ペアは一度のバンチ衝突で大量に生成されるので、少ない衝突で十分な データ量を得られるため次のトレインにフィードバックをかけることも可能である。

本章ではペアモニターの測定原理とペアモニターデザインについて述べる。

3.2 測定原理

ビーム中に含まれる粒子は対向するビームによる電磁場によって制動放射を起こす。この制動放射はビームストラールングと呼ばれ、電子・陽電子散乱による仮想光子とともに電子陽電子対生成の原因となる。この電子陽電子対生成は主に3つのインコヒーレント過程に大別される。それらのファインマン図を図3.1に示す。

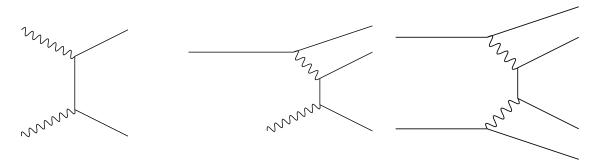


図 3.1: 3 つの Incoherent 過程。左図は Breit-Wheeler, 中央は Bethe-Heitler, そして右図は Landau-Lifshitz 過程と呼ばれる。それぞれの過程が繰り返し発生することによって低エネルギーな電子・陽電子ペアが大量に生成される。

これらの過程によって生成された粒子 (今後は Incoherent Pair と呼ぶ) は対向するビームと同電荷の粒子と異電荷の粒子に分けられる。それぞれの粒子を"同電荷"粒子と"異電荷"粒子と定義すると、異電荷粒子は対向するビームによる引力によってビーム中を振動しながら進む。同電荷粒子は対向するビームの強い電磁場から斥力を受けて散乱される。図 3.2 にこれらの現象の様子を示す。

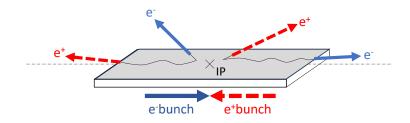


図 3.2: 電子陽電子対生成の様子 [5]。平坦な電子ビームと陽電子ビームが衝突点 (IP) で衝突し電子・陽電子ペアが生成されている様子を示している。異電荷粒子はビーム面を振動し、同電荷粒子は大きく散乱される。

この一連の現象はビームから発生する 2 次元のクーロンポテンシャルによる散乱として説明できる。ビームよる電磁場はビーム静止系から重心系にローレンツ変換すると、ビームの進行方向の電場成分はローレンツ収縮によって減少する。したがって、電磁場の成分は鉛直成分と水平成分の 2 次元のみになる。重心系で電場の垂直成分を E_{\perp} とすると次のように表される。

$$\boldsymbol{E}_{\perp} = \gamma \boldsymbol{E}_{\perp}^{\prime} \tag{3.3}$$

ここで、 E'_{\perp} はビーム静止系の電場の垂直成分である。同様に、重心系における磁場の鉛直成分は

$$\boldsymbol{B}_{\perp} = -\gamma \boldsymbol{\beta} \boldsymbol{E}_{\perp}^{\prime} \tag{3.4}$$

と表される。ここで β はビーム静止系に対する重心系の速度である。この電磁場中で電子・陽電子ペアが受けるローレンツカFは、電荷e、ビームの速度 β_{beam} とすると、

$$F = e(E_{\perp} + \beta_{beam} \times B_{\perp}) \tag{3.5}$$

で示される。ここで、陽電子ビームと同じ進行方向に進む粒子について考える。このとき 陽電子ビーム静止系に対する重心系の速度は $\boldsymbol{\beta}=(0,0,\beta)$ と仮定する。また、ビームの 速度は電子ビームを $\boldsymbol{\beta_{beam}}=(0,0,\beta_e)$ 、陽電子ビームを $\boldsymbol{\beta_{beam}}=(0,0,-\beta_p)$ とする。こ こで、電子ビームの磁場成分は

$$\beta_{beam} \times \boldsymbol{B}_{\perp} = \beta_{beam} \times (-\beta \times \boldsymbol{E}_{\perp})$$

$$= (\beta \beta_e E_x, \beta \beta_e E_y, 0)$$

$$= \beta \beta_e \boldsymbol{E}_{\perp}$$
(3.6)

と表される。したがって、粒子が電子ビームから受ける力は

$$\mathbf{F} = e(1 + \beta \beta_e) \mathbf{E}_{\perp} \tag{3.7}$$

で示される。同様に陽電子ビームから受ける力は

$$\mathbf{F} = e(1 - \beta \beta_p) \mathbf{E}_{\perp} \tag{3.8}$$

となる。 $\beta \simeq \beta_e \simeq \beta_p \simeq 1$ であるため、粒子は進行方向が同じである陽電子ビームからの影響は受けず、対向する電子ビームの影響のみを受ける。電子ビームと同じ進行方向に進む粒子について考えても、電子ビーム静止系に対する重心系の速度が $\beta = (0, 0, -\beta)$ となることから、結局粒子は電子ビームから影響を受けず陽電子ビームからの影響を受ける。

以上より生成された粒子は対向するビームによる影響を受ける。ビームの形状が変われば生成された粒子の散乱具合も変化する。ゆえにペアモニターで散乱具合を測定できれば対向するビームの情報を得ることができる。

第4章 シミュレーション設定

4.1 ビームパラメータ

本研究はビーム衝突シミュレーションソフトウェアを使用して Incoherent Pair を生成する。シミュレーションの座標を図 4.1 のように定義する。

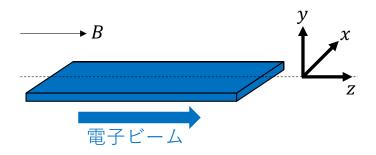


図 4.1: 座標

電子ビームが進む方向をz軸とし、z軸に対して鉛直方向をy軸、両方の軸に対して垂直な軸をx軸とする。シミュレーションを行うためにはビームパラメータの情報を定義する必要があり、本研究は Technical Design Report[1] に記載されている表 4.1 に示すパラメータを用いた。

表 4.1: TDR2013 ビームパラメータ

₹ 4.1. 1D1(2013 C	4//	<u> </u>	
Centre-of-mass energy	E_{CM}	250	GeV
Bunch population	N	2.0	$\times 10^{10}$
RMS bunch length	σ_z	0.3	mm
Horizontal emittance	$\gamma \epsilon_x$	10	$\mu\mathrm{m}$
Vertical emittance	$\gamma \epsilon_y$	35	nm
IP horizontal beta function	β_x^*	13.0	mm
IP vertical beta function	β_y^*	0.41	mm
IP RMS horizontal beam size	σ_x^*	729.0	nm
IP RMS vertical beam size	σ_y^*	7.7	nm

本研究はビームサイズ $\sigma_{x(y)}$ のほかに図 4.2 に示す衝突点から鉛直方向へのずれ δ_y も考慮する。 δ_y は対向するビームの鉛直方向のビームサイズに対してどれだけ相対的に離れ

ているかを表し、 $\delta_y \times \sigma_y$ と定義する。垂直方向のずれ δ_y の値がゼロのとき理想的なビーム軸上で衝突したことを意味し、値が大きくなるにしたがってビームが衝突点から遠ざかるように変位する。

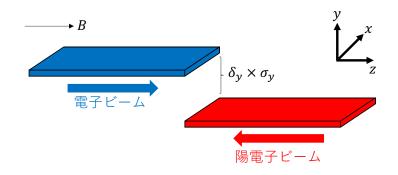


図 4.2: 鉛直方向のずれ δ_y

4.2 ビーム衝突シミュレーション

ビーム衝突シミュレーションを行うために CAIN と呼ばれるソフトウェアを用いた [6]。

4.2.1 CAIN

CAIN は高エネルギーの電子、陽電子そして光子による相互作用をシミュレーションするためのソフトウェアである。CAIN の前身となるソフトウェアは ABEL[7] と呼ばれ、1985 年頃に開発された。ABEL は電子・陽電子線形加速器のビームビーム相互作用のシミュレーションを行うことができる。後に、対生成による効果や放射光加速器シミュレーションも取り入れられ CAIN と改名した。本研究は CAIN 24.2 を使用する。

4.2.2 Incoherent Pair のトラッキング

ILD内は Z 軸の正の方向に 3.5T のソレノイド磁場がかかっている。加えて X 軸方向に anti-DID (Detector Integrated Dipole) という磁場がかかっている。衝突点でビームストラールングにより発生した電子・陽電子ペアは検出器に衝突することでノイズの原因になる。anti-DID は低エネルギーの Incoherent Pair を BeamDump へ繋がるパイプ方向に導き、他の検出器への影響を減少させる役割がある [1]。本研究はソレノイド磁場のみを考慮し anti-DID による磁場は考えない。また、衝突点からペアモニターまでの間に存在する他の検出器は考慮しない。

4.2.3 シミュレーションの結果

CAIN ではシミュレーションで使用する位置を t で表現しているため、以下の記述もその表現に従う。シミュレーションは電子ビームの中心が衝突する前の位置 t_0 から十分に相手のビームによる影響が無くなる位置 t_f まで行った。本研究は $t_0=-1.25mm$ から $t_f=1.25mm$ までとした。また衝突前から衝突後までの距離を n_t 個に分割し、電子ビームが進んだ距離を t_i を用いて表す。

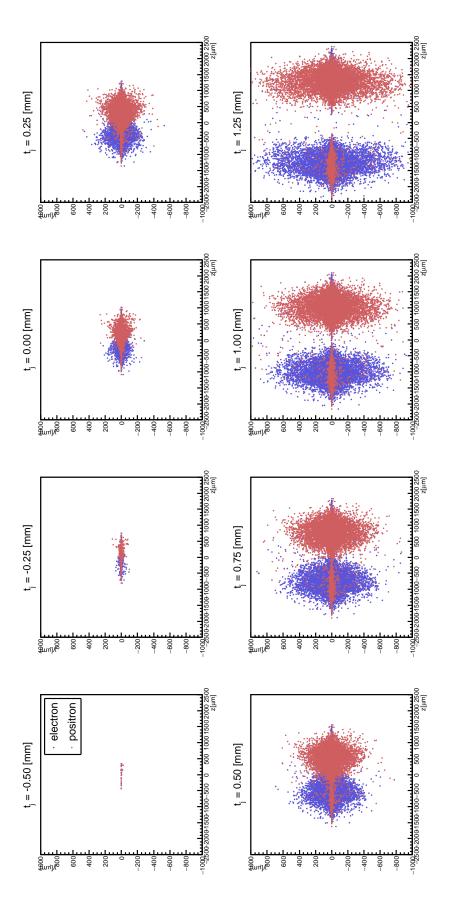
$$t_j = t_0 + j\Delta t \tag{4.1}$$

ここでjは n_t 分割の何番目であるかを表し、 Δt は

$$\Delta t = (t_f - t_0)/n_t \tag{4.2}$$

とする。図 4.3 に生成された Incoherent Pair の時間発展の様子を示す。ただし、 $t_j=-0.50$ mm までは Incoherent Pair が生成されなかったため図中に示していない。図中の青点は電子の位置を表す。同様に赤点は陽電子の位置を表す。 t_j は電子ビームの中心位置を示し、 $t_j=0.0$ mm のとき電子ビームの中心が原点にくる。時間発展にしたがって同電荷粒子の散乱する様子が見て取れる。

シミュレーション直後の+z方向から衝突点 IP 方向を見たとき、+4 m に設置したペアモニターへ向かう Incoherent Pair の位置を図 4.4 に、中心付近を拡大した分布を図 4.5 に示す。図 4.4 よりビームシミュレーション後の電子・陽電子は分布を示している。縦軸は Y 方向、横軸は X 方向とする。どちらの粒子も陽電子ビームによる電場の影響を受けているため、電子には引力が働き中心付近に集まる。また陽電子には斥力が働き散乱される。図 4.5 より、異電荷粒子となる電子はビーム面を振動するためビーム形状と同じように水平方向に広い分布となる (左)。同電荷粒子となる陽電子は斥力により上下に分布する (右)。



青 (赤) 点が電子 (陽電子) の位置を表す。 t_j は電子ビームの中心位置を示し、 $t_j=0.0~\mathrm{mm}$ 図 4.3: Incoherent Pair 生成の時間発展。 のとき電子ビームの中心が原点にくる。

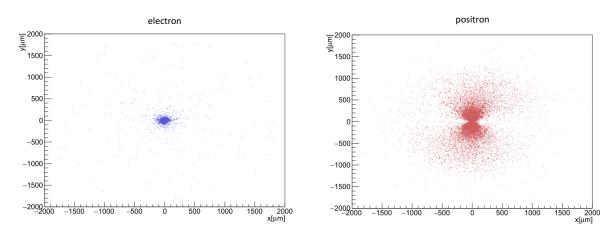


図 4.4: Incoherent Pair 分布。どちらの粒子も陽電子ビームによる電場の影響を受けているため、電子には引力が働き中心付近に集まる。また陽電子には斥力が働き散乱される。

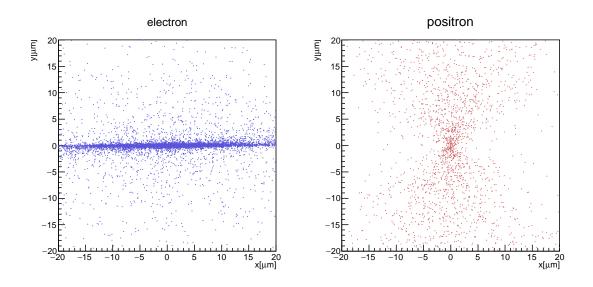


図 4.5: Incoherent Pair 分布 (拡大図)

第5章 解析

5.1 nominal なパラメータの場合の解析

5.1.1 Incoherent Pair の生成

ペアモニターは ILD 検出器内に 2 つ設置される。それらは ILD 検出器内の両端に設置され、第 3 章の原理から一方のペアモニターは電子ビームの情報を読み取り、もう一方のペアモニターは陽電子ビームの情報を読み取る。どちらのビームの情報を読み取り議論を進めようとしているのか明確にするため、今後の章で議論するペアモニターは図 5.1 に示すように陽電子ビームが衝突点へ向かう側の 4m の距離に設置されたペアモニターとする。 つまりそこから得られるヒット分布からは陽電子ビームの情報が得られる。

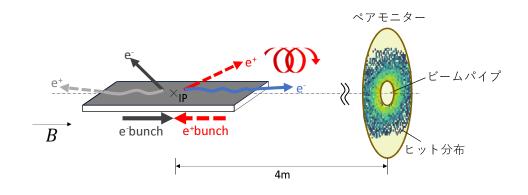


図 5.1: ペアモニターの位置

TDR[1] に記載されている設計値のもとビームシミュレーションを行い、図 5.2 の位置 からみた Incoherent Pair の位置を見た。そこから運動量 p の粒子が x-y 平面上で x 軸とな す角度を φ として Incoherent Pair の φ 方向の分布を求めた。図 5.3 にその φ 分布を示す。 $\varphi=0,\pm\pi\mathrm{rad}$ のとき x 軸上を向いている。

5.3 より Incoherent Pair の電子はビームから上下方向の引力を受けてビーム平面上を振動するように進行するため、x 軸上の $\varphi=0,\pm\pi\mathrm{rad}$ にピークが現れた (図 5.3 の左側)。対して、ビームと同電荷である陽電子はビームからの斥力を受けて $\varphi=\pm\pi/2\mathrm{rad}$ にピークが現れた (図 5.3 の右側)。

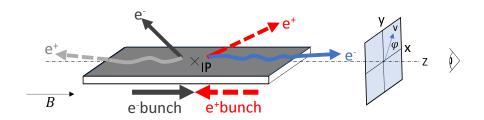


図 5.2: Incoherent Pair を見る位置の定義

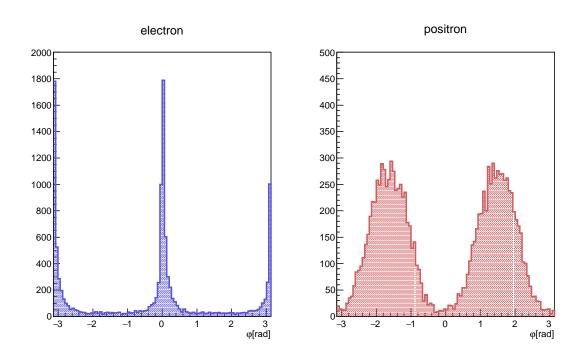


図 5.3: φ 分布。縦軸はエントリー数、横軸は x 軸に対する角度で $-\pi$ rad から $+\pi$ rad までとする。 $\varphi=0,\pm\pi$ rad のとき x 軸上を向いている。電子はビームに沿って振動するため $\varphi=0,\pm\pi$ rad にピークができる。また陽電子は上下に散乱されるため $\varphi=\pm\pi/2$ rad にピークができる。

5.1.2 Incoherent Pair の解析

Incoherent Pair を 4m 先のペアモニターにヒットするまでの運動の計算を行った。図 5.1 に示すペアモニターの位置まで 3.5T の一様なソレノイド磁場のもとに解析的に解いた。結果を図 5.4 に示す。

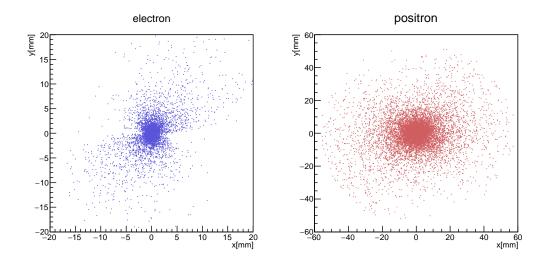


図 5.4: ペアモニター上での Incoherent Pair。左が電子のみのヒット分布、右が陽電子のみのヒット分布を示す。

図より電子はほぼ点対称な右上から左下にかけて広がる分布となり、陽電子は電子と比べてほぼ一様な分布となった。このような分布になった理由を考察する。

一様な磁場中で粒子の回転半径rは次の式から求められる。

$$r[m] = \frac{p_t[GeV/c]}{0.3B[T]} \tag{5.1}$$

ここで、 p_t は Incoherent Pair が持つ φ 方向の運動量 p_t 、B は磁場の大きさを表す。磁場の大きさは本研究の条件より一定であるため、回転半径は p_t に依存する。

まず電子に注目して、図 5.4 から電子が持つ p_t を $0.0\sim 2.0$ MeV/c, $2.0\sim 4.0$ MeV/c, ... と一定の範囲で分割し、それぞれの電子分布を図 5.5 に作成した。

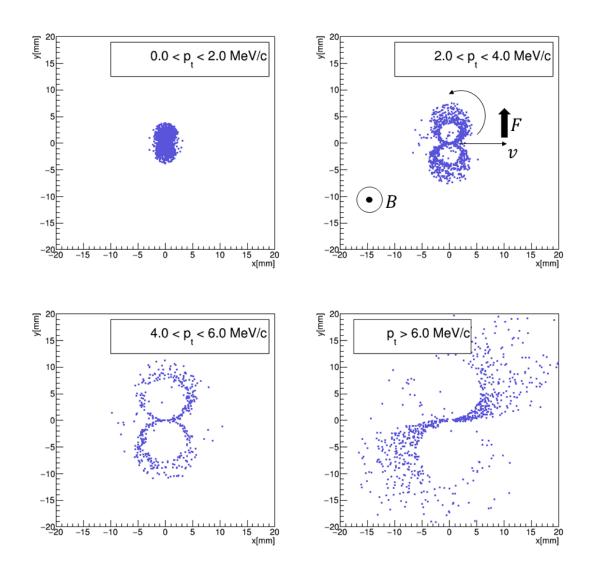


図 5.5: Incoherent Pair(電子)を解析的に解いた結果 $(p_t \, \mathbb{I})$

図より、 p_t が小さい電子はグラフの中心付近に分布し、 p_t が大きくなると分布が広がりながら円を描くように分布することがわかった。分布の広がりは式 5.1 から求められる半径によって求めることができ、円は磁場とローレンツ力の関係を用いて図 5.3 で電子の p_t が $\varphi=0$, $\pi {\rm rad}$ 方向を向いていることから上下に描かれる。また、 p_t の大きさによって、円の形状が 8 のような形状から乙字のような形状に変化する。これは p_t が大きい粒子が 1 周する前にペアモニターへ衝突したためだと考えられる。

分布図を分割したときと同様の範囲で、電子がペアモニターに衝突するまでの回転数を 調べた。その結果を図 5.6 に示す。

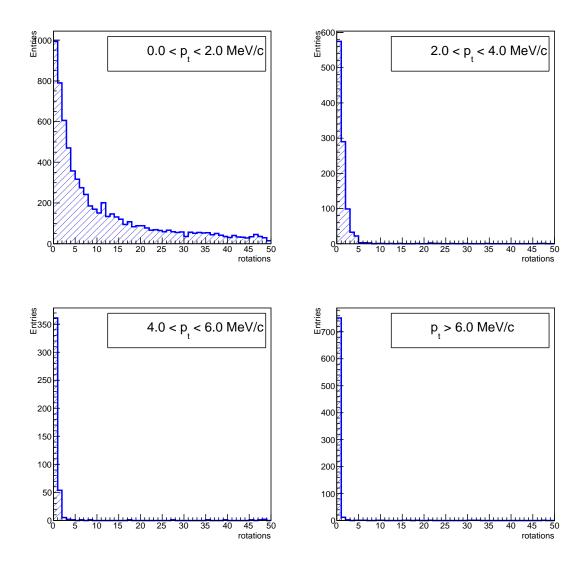


図 5.6: 電子がペアモニターに衝突するまでの回転数

図より p_t が小さいときは回転数が少ない粒子もあれば回転数が多い粒子もある。しかし、 p_t が大きくなると回転数が多い粒子は少なくなり、1 回転もせずにペアモニターに衝突する粒子が増える。これは p_t が高い粒子はビーム軸を向く運動量 p_z の値も高いためだと予想される。

以上のことから $0.0 < p_t < 2.0 \ {
m MeV/c}$ の電子は回転半径が小さくペアモニターに衝突するまで多く回転する粒子も存在するため図の中心付近で一様に分布する。 $2.0 < p_t < 6.0 \ {
m MeV/c}$ の電子は回転半径が大きくなり、衝突するまでの回転数も減少し、 $p_t > 6.0 \ {
m MeV/c}$ の電子は 1 回転する前にペアモニターに衝突するため図のような分布となった。

陽電子についても同様に p_t を一定の範囲で分割する。分布図を図 5.7 に、 p_t と回転数の関係を図 5.8 に示す。

図 5.3 より陽電子のピークが $\varphi = \pm \pi/2$ rad であることからローレンツ力は水平方向を向く。そのため陽電子は左右に円を描くように分布したと考えられる。しかし、その分布

は電子と比べて一様に分布する。これは陽電子の分布は φ 分布が $\pm\pi/2$ rad を中心として幅を持つ分布であることからもわかるように、陽電子の p_t の向きが様々な方向を向いているためである。

第 4 章で述べたようにビームサイズが $\sigma_x \gg \sigma_y$ であることが電子と陽電子で φ 分布に違いを生んだ。

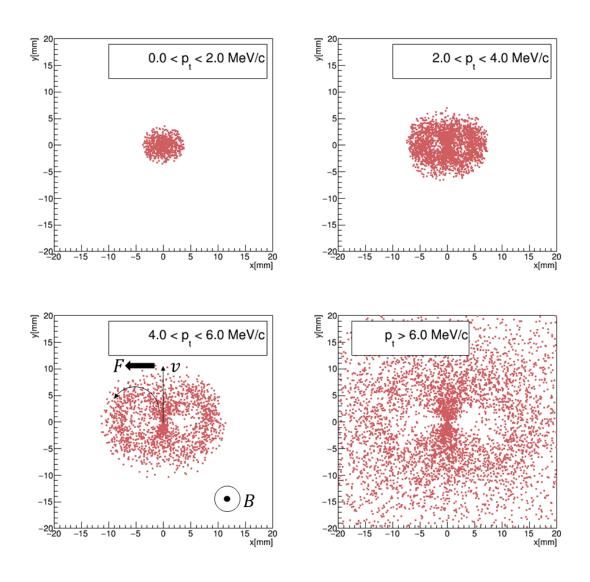


図 5.7: Incoherent Pair(陽電子) を解析的に解いた結果 $(p_t \mathbb{N})$

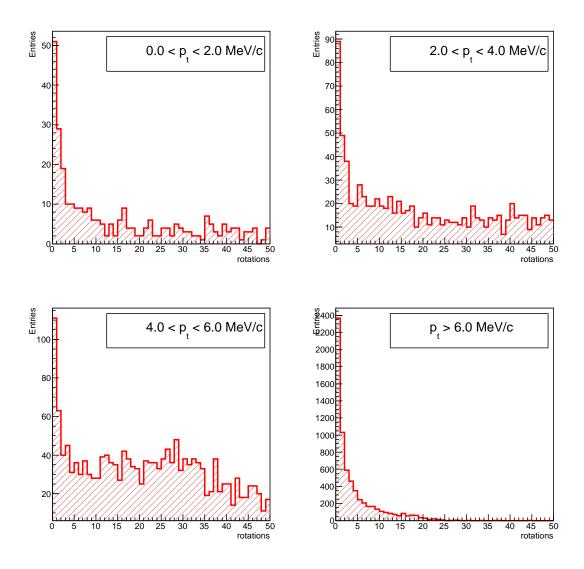


図 5.8: 陽電子がペアモニターに衝突するまでの回転数

5.1.3 ペアモニター上でのヒット分布

次にビームシミュレーションによって生成された Incoherent Pair を用いてペアモニターにヒットした粒子の解析を行う。第 5.1.1 節で述べたように、本論文では陽電子が大きく散乱される側のペアモニターの結果を示すため、ペアモニターのヒット分布を形成する要因となるのは陽電子になると予想される。ペアモニターにヒットした粒子数の分布を図5.9 に示す。また、ペアモニターにヒットした粒子の中で電子と陽電子をそれぞれ分けてプロットした散布図を図5.10 に示す。電子はほとんどがビームパイプ内を通過するのに対して陽電子はビームパイプ外にも広く分布している。つまり図5.9 中のヒットした粒子はほとんどが陽電子である。

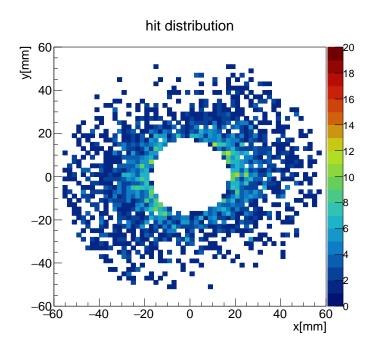


図 5.9: ペアモニターヒット分布図 (nominal パラメータ)。カラーの変化はヒットした粒子数を表す。

5.2 ペアモニター上ヒット分布の σ_x 依存性

ビームパラメータを nominal な値から変更したときペアモニターのヒット分布に変化が無ければ機械学習で設定したネットワークがビームパラメータを変更したときに得られる特徴を捉えることができず学習を行うことができない。そこで、 σ_x の値を変えて Incoherent Pair の生成およびペアモニターのヒット分布解析を行い、分布が変化するかどうか検討を行った。さらにこの設定で得られた結果と nominal な値の結果と比較した。今回はビームサイズの nominal な値を $\sigma_x^{TDR} = 729.0$ nm として、 $\sigma_x = 1.5\sigma_x^{TDR}$ 、 $\sigma_x = 3.0\sigma_x^{TDR}$ 、そして $\sigma_x = 5.0\sigma_x^{TDR}$ とサイズを変えて解析を行った。

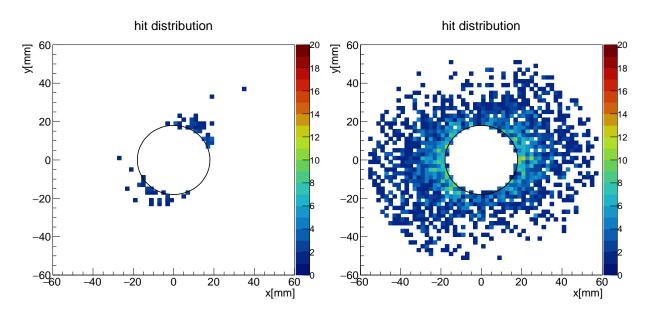


図 5.10: ペアモニターヒット分布図 (電子陽電子別)。図中の円はビームパイプの外形を表す。左図が電子のみのヒット分布図、右図が陽電子のみのヒット分布図である。

対向ビームが作る電場は単位面積あたりの電荷密度 ρ に比例する。ビーム形状は非常に平坦 $(\sigma_x \gg \sigma_y)$ なので、 σ_x の値を倍にすれば平面ビームの面積が倍になることから、単位面積あたりの電荷密度が減るため対向ビームによる電場が弱くなる。したがってビーム中の粒子が対向ビームによる電場によって Incoherent Pair を生成される機会が減り、Incoherent Pair の総数が減少すると考えられる。

5.2.1 Incoherent Pair の解析

シミュレーション直後の+z 方向から衝突点 IP 方向を見たときの Incoherent Pair の位置をを図 5.11 に示す。また、Incoherent Pair の φ 分布を図 5.12 に示す。

図 5.12 より、 σ_x の値が大きくなると電場が弱くなるため生成される粒子数が減少する。

5.2.2 nominal な場合との比較

前節で生成した Incoherent Pair をペアモニターの位置まで解析的に解いた。その結果を図 5.13 に示す。

また、 σ_x をさらに大きくした場合、例えば $\sigma_x = 5.0\sigma_x^{TDR}$ とした場合のペアモニターヒット分布図を作成し比較した。結果を図 5.14 に示す。 $\sigma_x = 5.0\sigma_x^{TDR}$ となると粒子がほぼ生成されなくなることが分かる。

図 5.13 からビームパイプより大きく散乱される粒子が図 5.4 と比べて減っていることが分かる。ヒット分布の広がりがどのように減少したか評価するため、座標の中心からペアモニターにヒットした粒子の総数の 97.5%を含む半径を最大半径 R_{max} と定義し、nominal な場合の結果と $\sigma_x=1.5\sigma_x^{TDR}$ 、 $\sigma_x=3.0\sigma_x^{TDR}$ としたときの結果を図 5.15 に示す。 図よ

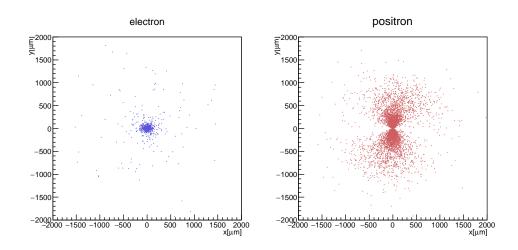


図 5.11: シミュレーションによって生成された Incoherent Pair $(\sigma_x = 1.5\sigma_x^{TDR})$ 。

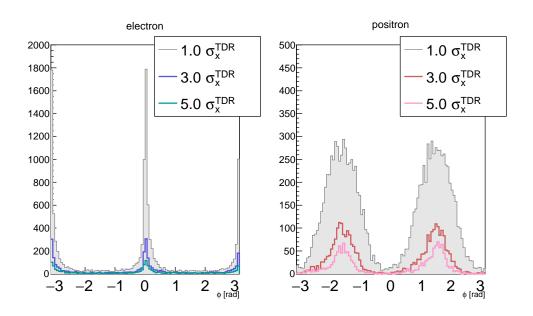


図 5.12: φ 分布の σ_x 比較。縦軸はエントリー数。灰色で表示したヒストグラムは nominal な場合の結果を示していて、その他のヒストグラムは $\sigma_x=3.0\sigma_x^{TDR}$ および $\sigma_x=5.0\sigma_x^{TDR}$ としたときの結果である。

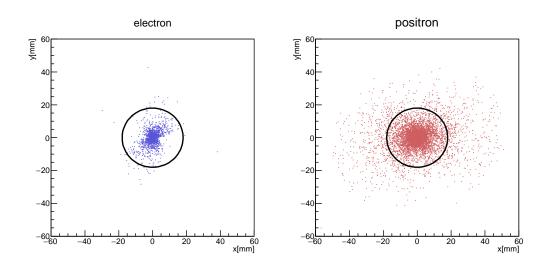


図 5.13: ペアモニターヒット分布図 $(\sigma_x=1.5\sigma_x^{TDR}$ 、電子陽電子別)。左が電子のみのヒット分布、右が陽電子のみのヒット分布を示す。中央の円はビームのパイプの外形を表している。直径は 18.0mm とする。

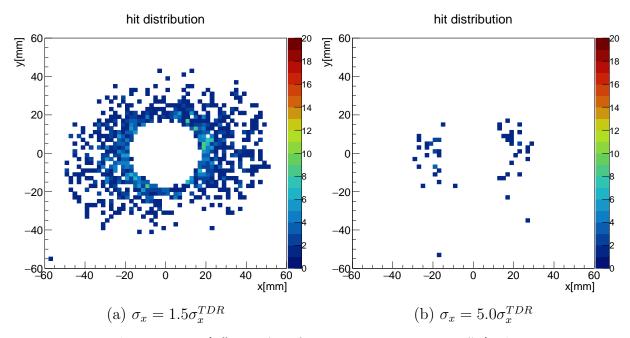


図 5.14: σ_x を変化させたときのペアモニターヒット分布図

り σ_x を変化させるとビームの中心から離れるほど粒子が多く減少し、最大半径 R_{max} が小さくなることがわかった。これは、ビームによる電場が低くなったため散乱される粒子が減少したからだと考えられる。

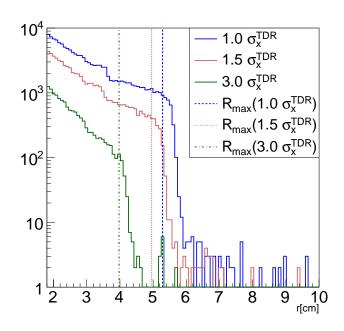


図 5.15: 動径方向のヒット分布。縦軸がエントリー数。データは 50 バンチ分。青線が水平方向のビームサイズ σ_x が nominal な場合として、赤線がその値から 1.5 倍にした場合、緑線のヒストグラムが 3.0 倍にした場合の結果である。点線がそれぞれの最大半径 R_{max} を表す。

5.3 ペアモニター上ヒット分布の σ_y 依存性

前節と同様に σ_y に変更して解析を行った。ビームサイズの nominal な値 $\sigma_y^{TDR}=7.7$ nm として、 $\sigma_y=1.5\sigma_y^{TDR}$ 、 $\sigma_y=3.0\sigma_y^{TDR}$ 、そして $\sigma_y=5.0\sigma_y^{TDR}$ とビームサイズを変えた。このように一方のビームの鉛直方向のビームサイズが対向するビームのサイズよりも小さいとき、ビーム中の粒子同士が相互作用する割合が少なくなってしまうため、衝突によって生成される Incoherent Pair の総数が減少すると予想される。

5.3.1 Incoherent Pair の解析

シミュレーション後の+z 方向から衝突点 IP 方向を見たときの Incoherent Pair の位置を図 5.16 に示す。図 5.17 その φ 分布を示す。図 5.17 より、生成される粒子数は減少するが、 σ_y の値が大きくなったことで電場の水平方向の成分による影響が出るようになり、減少の度合いは σ_x と異なる。

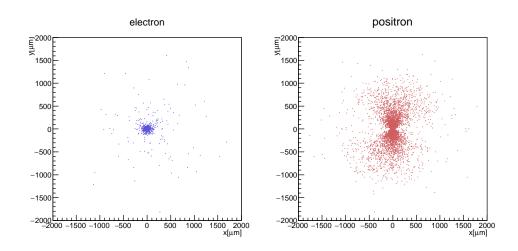


図 5.16: シミュレーションによって生成された Incoherent Pair($\sigma_y=3.0~\sigma_y^{TDR}$)。

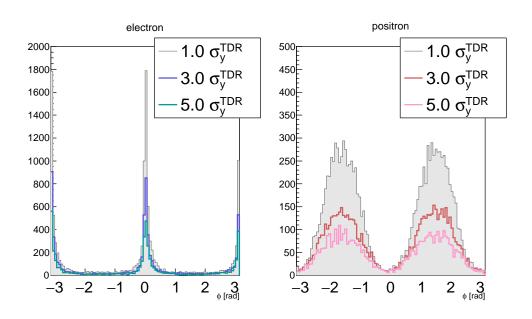


図 5.17: φ 分布の σ_y 比較。縦軸はエントリー数。灰色で表示したヒストグラムは nominal な場合の結果を示していて、その他のヒストグラムは $\sigma_x=3.0\sigma_x^{TDR}$ および $\sigma_x=5.0\sigma_x^{TDR}$ としたときの結果である。

5.3.2 nominal な場合との比較

前節で生成した Incoherent Pair がペアモニターにヒットした痕跡を残すとき、水平方向の成分を含む粒子の割合が増えているため、ヒット分布は nominal な場合と比べてより一様な分布になると考えられる。 $\sigma_y=1.5\sigma_y^{TDR}$ の場合と $\sigma_y=5.0\sigma_y^{TDR}$ の場合でそれぞれ Incoherent Pair を生成させ、ペアモニターの位置まで解析的に解いた。ペアモニターのヒット分布を図 5.18 に示す。図よりヒットした粒子は減少したが、その減少の度合いは σ_x を変更した場合と比べて少ない事がわかった。

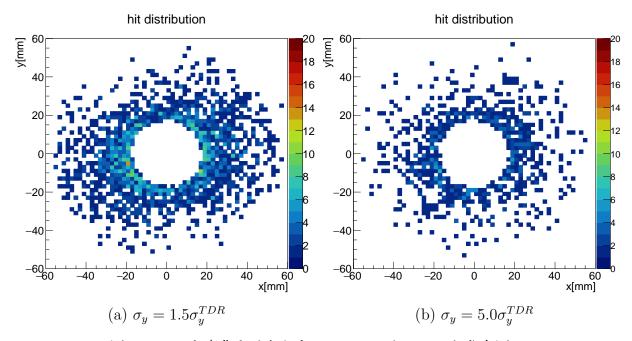


図 $5.18: \sigma_y$ を変化させたときのペアモニターヒット分布図

どのように減少したか確かめるため、最大半径 R_{max} を計算をした。ペアモニターのヒット分布から最大半径 R_{max} を求めて、nominal な場合と $\sigma_y=1.5\sigma_y^{TDR}$ と $\sigma_y=3.0\sigma_y^{TDR}$ の場合と比較を行った。その結果を図 5.19 に示す。図 5.15 と比べて変化は少なく、最大半径は σ_x の変化に支配的であることがわかった。

5.4 先行研究との比較

先行研究 [8] では ILC 設計値を用いたシミュレーションによってペアモニターにヒットする電子・陽電子の分布を再現しビームサイズの測定方法の研究を行った。この ILC の設計値は 2007 年に発行された Reference Design Report(RDR)[9] に記述されている設計値に基づいている。研究結果よりビームサイズ (σ_x,σ_y) とビーム同士のずれ δ_y に関して ILC で要求されている精度 (10%) を満たす再構成精度が得られた。その後、TDR が発行され ILC の設計値に変更が加えられたため、先行研究 [10] で TDR の設計値についても再構成が可能かどうか検証した。ペアモニターのヒット分布は利用せず、それぞれの設計値で生成された Incoherent Pair の粒子の類似性を検証することで、RDR の場合と同様に TDR

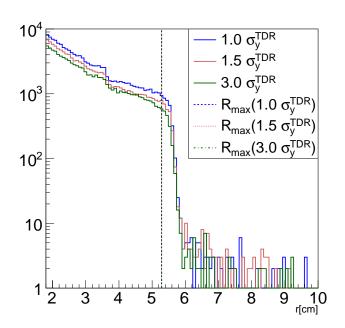


図 5.19: 動径方向のヒット分布。縦軸がエントリー数。データは 50 バンチ分。青線が鉛直方向のビームサイズ σ_y が nominal な場合の結果、赤線がその値から 1.5 倍にした場合の結果そして緑線が 3.0 倍にした場合の結果である。点線がそれぞれの最大半径 R_{max} を表す。

設計値でも再構成が可能か研究を行った。その結果、粒子の総数について RDR と TDR で類似性が確かめられ、TDR の設計値を用いてペアモニターのヒット分布からビームサイズの再構成が行えることを示唆した。

先行研究 [8] は RDR のパラメータを用いてペアモニターのヒット分布の再現を行った。本研究の妥当性を確かめるため、先行研究のパラメータを用いてペアモニターのヒット分布を作成し比較検討する。先行研究で用いていたパラメータを以下の表 5.1 に示す。ペアモニターの位置は先行研究を参考に 4.275m の位置に設置したと仮定した。また、先行研究は物質との相互作用も考慮するフルシミュレーションを行い、ビームパイプ内を進行する粒子を最大化する anti-DID という磁場を加えている。しかし、本研究はその二つは考慮せず一様なソレノイド磁場のみを考慮してシミュレーションを行っている。また、陽電子ビームが通過するビームパイプも考えない。

図 5.20 にビームパラメータ RDR を用いて再現したペアモニターのヒット分布図、図 5.21 に先行研究のヒット分布図を示す。

両者の図を比較して、粒子の分布範囲がフルシミュレーションを行ったほうが解析的に解いた場合よりも大きいことがわかった。anti-DIDによってILD内に分布する磁場は次の図 5.22 のように変化する。ペアモニターはRDRの場合衝突点から 4.257m であるので、磁場の磁力線が広がる箇所に位置する。衝突後に生成された Incoherent Pair は磁力線に沿って進行するため、ペアモニターにヒットした粒子の分布が広がったのだと考えられる。

表 5.1: ビームパラメータ [8]。* は本研究と異なる値のパラメータを示している。

Parameter		RDR	
Centre-of-mass energy*	E_{CM}	500	GeV
Bunch population	N	2.0	$\times 10^{10}$
RMS bunch length	σ_z	0.3	mm
Horizontal emittance	$\gamma \epsilon_x$	10	$\mu\mathrm{m}$
Vertical emittance*	$\gamma \epsilon_y$	40	nm
IP horizontal beta function*	eta_x^*	20	mm
IP vertical beta function*	β_y^*	0.40	mm
IP RMS horizontal beam size*	σ_x^*	639	nm
IP RMS vertical beam size*	σ_y^*	5.7	nm

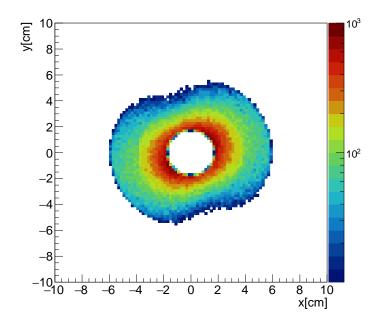


図 5.20: ペアモニターヒット分布図 (50 バンチ分)。

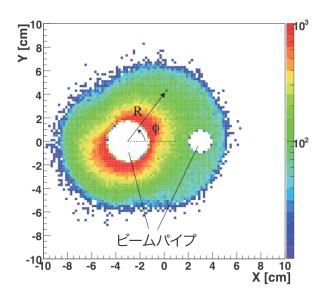


図 5.21: ペアモニターヒット分布図 [8]。データは 50 バンチ分。ビームパラメータは RDR の値を使用しフルシミュレーションを行った。左側の穴が電子の出射ビームパイプ (半径 $1.8 \mathrm{cm}$)、右側の穴が陽電子の入射ビームパイプ (半径 $1.0 \mathrm{cm}$)。

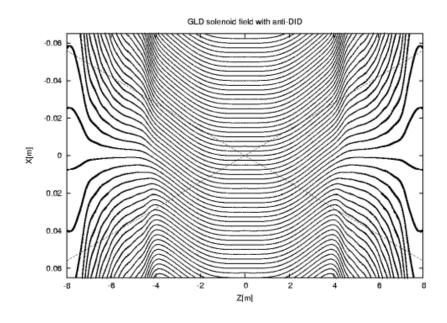


図 5.22: anti-DID による効果。anti-DID によって磁力線はビープパイプを沿うように分布する。anti-DID の強さはビームパイプ内を進行する Incoherent Pair が最大になるように最適化されている [8]。

第6章 機械学習

6.1 概要

前章で粒子は対向するビームの情報を持つことを述べた。ペアモニターはその粒子が残したヒット分布からビームサイズやビーム同士のズレといったビームパラメータを測定する。先行研究では、ヒット分布から求められる総ヒット数やヒット分布の特定の領域を測定変数と定義し、基準となるビームパラメータ付近で測定変数を用いてテイラー展開することでビームパラメータを求めた。本研究はヒット分布の全領域を使用する。その場合ヒット分布は画像のように捉えられる。これはビームパラメータを求めるとき測定変数を定義せずに済み、情報の欠如がない利点を持つ。

近年、大量の画像を入力として利用し、複数の物体の認識や分類する技術として機械学 習が注目されている。本章は機械学習の概要と今回使用した手法について述べる。

6.2 機械学習

1940 年代、ENIAC(Electronic Numerical Integrator and Computer) を代表とする弾道計算を行うための電子計算機の開発が盛んに行われた。その流れの中で一つの計算に特化せず、より汎用性のある計算を行うアルゴリズムの開発が進められる。ここから発展した技術が今日幅広い分野で用いられている機械学習のルーツである。1951 年に Marvin Minsky と Dean Edmonds らによってニューロンと記憶の関係を記述するヘッブ則に基づいたニューロンネットワーク SNARC(Stochastic Neural Analog Reinforcement Calculator)が発明された。そして、1952 年に Arthur Samuel がチェッカーズプログラムを世界で初めて開発した [11]。これらは機械学習の先駆けとなる研究であった。Frank Rosenblatt は生物の脳神経細胞 (ニューロン) のモデル化したパーセプトロンを 1957 年に提案された [12]。パーセプトロンは現在の機械学習でも使われている基本的な手法である。当時のメディアによって話題となり、第一次ニューラルネットワークブームが引き起こされた。ブームと同時期にパターン認識で優れた性能を発揮する Support Vector Machine (SVM) も提案された [13]。

しばらくして機械学習業界は落ち着きを取り戻すが、1980年代から 1990年代にかけて第二次ニューラルネットワークブームが再び訪れる。1980年に福島邦彦がネオコグニトロン [14] の提案、LeCun らがこれを元に手書き文字認識に応用させた畳み込みニューラルネットワークを開発した。1990年代から 2000年にかけてデータから学習する統計的学習の研究が進められた。これがきっかけとなりデータ解析に機械学習が使用されるように

第 6 章 機械学習 6.2. 機械学習

なった。また、WWWが誕生し世界中からデータを集めることが容易になり、そして計算機自体の性能が向上する。

2012年に物体の認識率を競うコンペティション ImageNet Large Scale Visual Recognition Competition (ILSVRC) において畳み込みネットワークを用いた手法 AlexNet [15] がこれまでの結果より精度に大差をつけて優勝した。そして同年、Andrew Ng や Jeff Dean らの所属する Google Brain チームがラベル付けなしに猫の顔や人体の顔を認識することに成功した [16]。これらのことが世界中で注目され、第三次ニューラルネットワークブームが巻き起こった。その後、ニューラルネットワークは医療分野や素粒子実験分野など幅広い分野で応用が進められている。

6.2.1 パーセプトロン

パーセプトロンは生物の脳内神経細胞を模したアルゴリズムである。人工ニューロンとも呼ばれる。パーセプトロンは1つ以上の入力と1つの出力を持つ。図 6.1 に 2 つの入力を受け取り、1つの出力を計算するモデルを示す。この図の場合、 x_1, x_2 が入力、z が出力

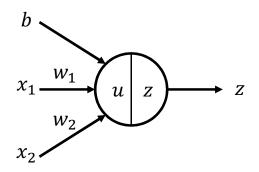


図 6.1: パーセプトロン概要図 (2 入力、1 出力)

を表している。入力にはそれぞれの重み w_1, w_2 が掛けられ、バイアスbを足した値が総入力となる。したがって、総入力uは

$$u = w_1 x_1 + w_2 x_2 + b (6.1)$$

で示される。出力zは総入力uを変数とする関数fによって求められる。

$$z = f(u) \tag{6.2}$$

この関数 f は活性化関数と呼ばれ、入力を出力へ変換する関数である。活性化関数については節 6.2.3 で詳しく述べる。

6.2.2 ニューラルネット

パーセプトロンは個数を自由に決めることができ、またある出力を別のパーセプトロンの入力として使用しても良い。ここで、図 6.2 のようなパーセプトロンを層状に並べた

6.2. 機械学習 第 6 章 機械学習

構造が2層あるネットワークを考える。このネットワークの計算は左から右へと進んでい

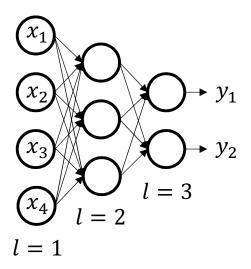


図 6.2: 2層構造のネットワーク。バイアスの表記は省略し、 x_n が入力、 y_n が出力を表す。 l=1 を入力層、l=2 を隠れ層、l=3 を出力層と呼ぶ。

く。左から順に各層を l=1,2,3 で表すと l=1 を入力層 (input layer)、 l=2 を隠れ層 (hidden layer)、 l=3 を出力層 (output layer) と呼ばれ、入力層・隠れ層・出力層で構成 されるネットワークをニューラルネットワークとする。また例で示した図 6.2 は隠れ層が 1 つしか無いがニューラルネットワークの隠れ層の数は自由に決めることができる。隠れ層が 2 層以上あるネットワークをディープニューラルネットワークと定義する。

6.2.3 活性化関数 (Activation)

活性化関数は神経細胞同時の伝達を模した関数である。活性化関数は非線形関数で入力の総入力を出力へどのように活性化するか決定する。活性化関数は様々な種類があるが次に示す関数が良く使用される。

ステップ関数

ステップ関数は出力を0か1で表現する活性化関数である。図6.3のように0を閾値として総入力が0より大きければ1を出力し、そうでなければ0を出力する。

$$f(x) = \begin{cases} 0 & (x \le 0) \\ 1 & (x > 0) \end{cases}$$
 (6.3)

第 6 章 機械学習 6.2. 機械学習

シグモイド関数

シグモイド関数は式 6.4 で定義される関数である。図 6.3 のように入力を与えると実数値を返し、そして出力はステップ関数と比べて連続的に変化する。

$$f(x) = \frac{1}{1 + e^{-x}} \tag{6.4}$$

ReLU(Rectified Linear Unit)

図 6.3 のように ReLU は入力が 0 より大きければ、その入力の値を出力し、0 以下であれば 0 を出力する活性化関数である。ReLU は式 6.5 のように表される。

$$f(x) = \begin{cases} x & (x > 0) \\ 0 & (x \le 0) \end{cases}$$
 (6.5)

ReLU は入力 x の x < 0 の部分を 0 に置き換えただけであるため計算量が小さい。これにより他の関数よりも早く学習を終了させることができ、ネットワークによっては上記の活性化関数よりも良い結果が得られることもある。そのため、近年は多くのネットワークで採用されている。

ソフトマックス (softmax)

ソフトマックスは入力xに対してK個のクラスに分類する問題の場合、例えば、手書き文字認識といった問題を解くときに使用される活性化関数である。クラス分類問題は出力層でK個のクラスのどれかに出力する。出力を y_1,\ldots,y_K と定義すると、出力層のk番目の出力は

$$y_k \equiv z_k = \frac{e^{u_k}}{\sum_{i=1}^K e^{u_i}} \tag{6.6}$$

によって示される。このとき、出力の総和は1となる。

$$\sum_{j=1}^{K} y_j = 1$$

6.2.4 損失関数 (Loss function)

ニューラルネットワークはバイアスを除いて基本的に全ての入力に重みwが掛けられる。つまり、wを変化させるとニューラルネットワークに影響を及ぼすため、ニューラルネットワークの構成で最も重要視される。ニューラルネットワークの全ての重みを成分に持つベクトルwを定義し、wを変化させることで、ニューラルネットワークが望んだ出力を計算するようにすることを考える。

6.2. 機械学習 第 6 章 機械学習

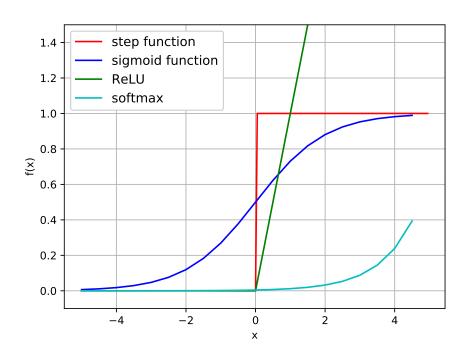


図 6.3: 活性化関数

入力の全ての成分をx、求めたい出力を目標値dと定義すると、入出力のペアは(x,d)と書くことでき、そしてそのペアが複数与えられているとする。

$$(x_1, d_1), (x_2, d_2), (x_3, d_3), \dots, (x_N, d_N)$$

この個々のペアは画像や音声データなど何を機械学習で学習させたいかに応じて異なる。 ここで、ペアの集合をトレーニングデータ (training data) と定義する。

今、あるペアの入出力をニューラルネットワークの入力と出力で再現できないか考える。すなわち、ニューラルネットワークにどのペアに対しても入力 x_n を与えた時、w を適当な値にすることで、そのニューラルネットワークの出力値 y_n が d_n とほぼ同様な値を取るようにする。この w を調整することを学習と呼ぶ。このとき、ニューラルネットワークによる出力とトレーニングデータの真値がどれだけ近い値なのか定量的に判断できる尺度が重要になる。この尺度を損失関数 (Loss function) と定義される。

損失関数 L は機械学習でどのような問題を解くのかによって形が異なる。クラス分類問題を解くときは次に示す交差エントロピー (cross entropy) と呼ばれる関数が使用される。

$$L = -\sum_{n=1}^{N} \sum_{k=1}^{K} d_{nk} \log y_{nk}$$
(6.7)

ここで、目標値 \mathbf{d}_n は 2 値の値を K 個並べたベクトル $\mathbf{d}_n = [d_{n1} \dots d_{nK}]^T$ で表現される。 \mathbf{d}_n の各値は、正解値に対する成分は 1、それ以外の成分は 0 となるように定義する。 例えば、 0 から 9 までの数値を認識する手書き文字認識問題で画像が示している数値が 8 のと

き、 d_n は

$$d_n = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0]^{\mathrm{T}}$$

で与えられる。この配列は最も左側の要素が手書き文字の0に、最も右側の要素が手書き文字の9に対応している。また、文字認識問題といったクラス分類のとき出力値 y_n は総和を1とする確率密度関数と同様な形で与えられる。例えば、上述の画像 (正解が8である画像) をニューラルネットワークに与えたとき、出力値は次式のような形式で表すことができる。

$$\mathbf{y}_n = [y_{n1} \ y_{n2} \ \dots \ y_{nK}]$$

= $[0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0.005 \ 0.985 \ 0.01]$

N=1と仮定したとき、損失関数は目標値が1の要素のみを計算すれば良い。 \mathbf{d}_n の要素が1のときの出力値は0.985であるため、対数をとると損失関数は0.015となる。

6.2.5 データセット

トレーニングデータによって調整されたネットワークが全く未知のデータに対しても正しく動作しなければならない。そこで、学習中にある段階で全く未知のデータを用いてネットワークの汎化性能を検証する。つまり未知のデータに対しても損失関数で計算を行い評価する。この未知のデータを検証データ (validation data) と呼ぶ。さらに、トレーニングデータや検証データとも異なるデータを用意し、学習後に最終的なネットワークの性能を評価する。このデータをテストデータ (test data) と呼ぶ。

6.3 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク (CNN; convolutional neural network) とは画像認識を行うときに良く用いられるネットワークの一種である。図 6.2 で表している層は隣接している層とすべての要素が結合していることから全結合層 (Full Connection) とも呼ばれる。CNN は全結合層に加えて畳み込み層とプーリング層の 2 つの層を含むネットワークを構成する。図 6.4 に畳み込みニューラルネットワークの例を示す。入力は画像のピクセル値であり、出力はラベルの正解率が出力される。

6.3.1 畳み込み層 (Convolution layer)

畳み込み層は画像に対して畳み込み演算を行う処理層である (図 6.6)。濃度の値を数値で表現した 1 枚の白黒画像を考える。画像サイズを $W \times W$ として、画像のピクセルを $(i,j)(i=0,\ldots,W-1,j=0,\ldots,W-1)$ で表す。また、ピクセルの数値を $x_{i,j}$ で表現し、実数値とする。ここで、フィルタと呼ばれる白黒画像より小さい画像を考える。フィルタのサイズは $F \times F$ として、フィルタのピクセルを $(p,q)(p=0,\ldots,F-1,q=0,\ldots,F-1)$

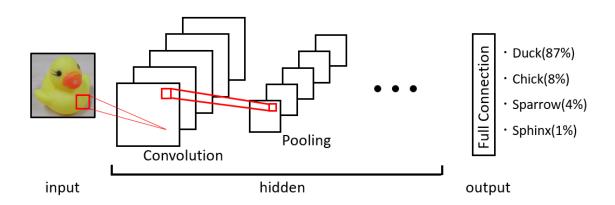


図 6.4: 畳み込みニューラルネットワークの例。入力が画像のピクセル値となる。隠れ層で畳み込み層とプーリング層が複数個呼び出すことで特徴を抽出する。出力はラベルの正解率となる。

で表す。フィルタのピクセル値は $m_{q,p}$ として、これも実数値とする。畳み込み演算は白 黒画像に対してフィルタを適応させる演算である。

$$u_{i,j} = \sum_{p=0}^{F-1} \sum_{q=0}^{F-1} x_{i+p,j+q} m_{p,q}$$
(6.8)

畳み込み演算は積和計算 (あるいは相関) と呼ばれる。この演算による結果をもとの画像 と対応する位置に格納し、すべてのピクセルで演算をする。

フィルタの枚数はユーザーが自由に設定することができる。本研究のニューラルネットワークではある畳込み層で 64 枚のフィルタを使用した。例として、その中の 60 枚のフィルタとそのフィルタを使用して畳み込み演算を行った出力結果を図 6.5 に示す。

畳み込み演算を行うことで出力画像サイズが縮小する。この縮小を避けるために画像の縁を 0 で埋めるゼロパディングと呼ばれる手法を用いた。

6.3.2 プーリング層 (Pooling layer)

プーリング層は畳み込み層の直後に実行される処理層で、画像サイズを小さくする演算処理を行う (図 6.7)。機械学習では \max プーリングと呼ばれる処理が良く使用される。例えば、図 6.8 のように画像サイズ $W\times W$ の画像で灰色の枠で示した $F\times F$ の領域を考える。この領域内で最も高い数値が出力される。この処理を画像の全領域で実行すると、その画像内で最も特徴がある要素が抽出される。

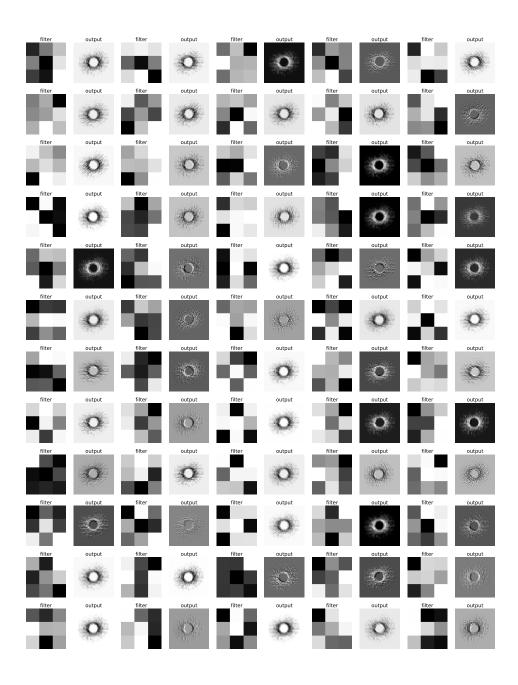


図 6.5: ある畳み込み層のフィルタとその出力結果。それぞれの画像は見やすさを考慮し実際のサイズと異なるサイズで表示している。3×3のサイズで表現された白黒画像がフィルタでその右隣にある画像がフィルタを適応させた後の出力結果である。

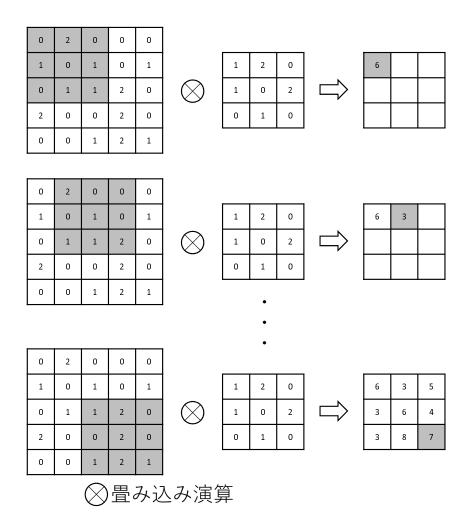


図 6.6: 畳み込み層の例。画像サイズは 5×5 として、フィルターサイズが 3×3 のとき、出力は 3×3 となる。

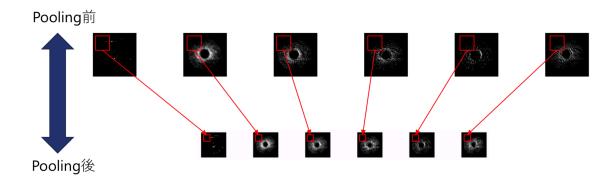


図 6.7: プーリング層概要。pooling 層によって赤枠で示したある領域のなかで最も特徴がある要素が抽出される。

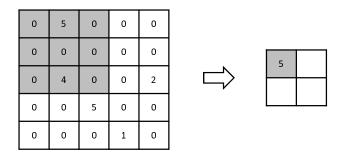


図 6.8: プーリング層の例。灰色の領域から最も特徴ある要素 (ここでは最も高い数値) だけを出力する。実行後画像サイズは小さくなる。

6.4 精度改善の手法

6.4.1 Dropout

Dropout はニューラルネットワークのパーセプトロンを確率的に仕分けて学習を行う手法である [17]。

その方法はある層で予め設定した確率 p に従ってランダムにパーセプトロンを選択しそれ以外を無効化する。無効化というのは初めからそのパーセプトロンが存在していなかったかのように扱う。例えば 4 つのパーセプトロンがある隠れ層で Dropout を適用すると確率 p=0.5 とした場合、図 6.9 のようになる。

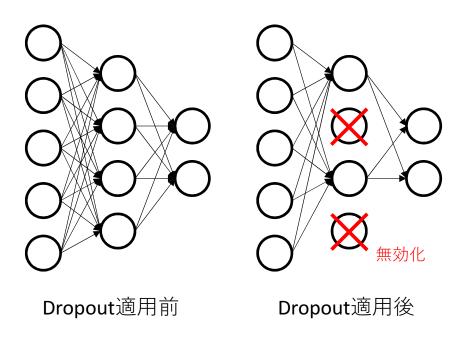


図 6.9: Dropout の例。隠れ層に Dropout を適応している。確率 p=0.5 でランダムにパーセプトロンを選択し、それ以外のパーセプトロンを無効化する。

そして、このネットワークの状態で学習を行い重みを調整する。パーセプトロンの選択は重みの更新するたびランダムに行う。Dropoutの目的は一定数のパーセプトロンをランダムに無効化させることでニューラルネットワークの自由度を下げて、ニューラルネットワークがトレーニングデータに対して過剰に適合してしまうことを防ぐことである。ただし評価時はパーセプトロンは無効化させず全て用いて評価を行う。このときパーセプトロンからの出力にpを掛ける。これは学習時のパーセプトロンの数が評価時の数に比べてp 倍になっていることを補償する。

6.4.2 Batch Normalization

Batch Normalization は入力値を平均 0、分散 1 に規格化させる [18]。

Batch Normalization と Dropout の併用を行うとニューラルネットワークのパフォーマンスが落ちることが報告されている。原因は Dropout は学習時と評価時でパーセプトロンの数が異なるため分散が変化するのに対して、Batch Normalization は学習時と評価時で分散を維持する。両者の分散が不一致であることによってパフォーマンスが低下する[19]。

本研究はBatch Normalization は使用せず、Dropout を Pooling 層の後に適用した。

6.5 使用ソフトウェア

機械学習は Tensoflow と Keras を用いた。Tensorflow は Google によって開発された機械学習を行うためのオープンソースライブラリ [20] である。また、Keras は Tensorflow の関数を扱いやすくするためのライブラリである。画像認識や音声認識、自然言語処理など分野で機械学習を行うとき、学習ネットワークが高度なプログラミング技術無しに実装できる強みを持つ。また、GPU ハードウェアのサポート、クラウド上での実行や学習状況の可視化など機械学習を効率化するツールが充実している。

第7章 機械学習への応用

7.1 ビームパラメータ (σ_y) の再構成

前章でペアモニターにヒットした粒子の分布形状やヒットした総数はビームパラメータに応じて変化することがわかった。このペアモニターヒット分布を入力として機械学習によってビームパラメータを再構成する手法を考える。初めに、多数あるビームパラメータの中から一つだけ選択しビームパラメータの再構成を試みる。その結果から機械学習を用いた再構成が妥当であるかどうか検証をする。ビームパラメータは鉛直方向のビームサイズ σ_y のみを変化させ、水平方向のビームサイズやビーム同士のずれは nominal な値とした。

7.1.1 トレーニングデータ

ビームサイズ σ_y は TDR 設計値を σ_y^{TDR} として一定間隔で変化させて、数パターン用意した。ビームサイズを変更させるのは陽電子ビームのみとし、電子ビームには変更を加えない。このビーム設定のもと CAIN によるビーム衝突シミュレーションを行い、生成された粒子をペアモニターまで解析的に解く。そして、ペアモニターにヒットした粒子の分布をグレースケールに変換し、この画像をトレーニングデータとして利用した。

ビーム衝突シミュレーションで使用する CAIN ソフトウェアは毎回異なる結果にすることが可能である。CAIN は擬似乱数を生成してシミュレーションを行う。この擬似乱数は CAIN 内部の"Rand"パラメータによって決定づけられる。この Rand パラメータはシード値と呼ばれ、シミュレーションを行う際に異なるシード値を与えることで出力される値が変化する。

7.1.2 3層ニューラルネットワークを用いた検討

初めに、機械学習が有効であるか検討するために σ_y を $1.0 \le \sigma_y/\sigma_y^{TDR} \le 10.0$ と $1.0\sigma_y$ ずつ変化させて学習を行った。各 σ_y に対して 500 枚の画像を用意し、合計 5000 枚の画像を含むデータセットを作成した。これらの画像をランダムにプログラムへ読み込ませ、最初の 9 割をトレーニングデータ、残りの 1 割を検証データとして利用した。また、このデータセットとは別に各 σ_y で 50 枚、合計 500 枚作成した (表 7.1)。このデータセットをテストデータとして使用する。

トレーニングデータ及びテストデータ用の画像は画像サイズを $64 \times 64 \times 8$ ビット256 階調でグレースケール化した画像とした。これはメモリ不足による学習失敗を防止する

σ_y/σ_y^{TDR}	トレーニング+検証 データ	テストデータ
1.0	500	50
2.0	500	50
3.0	500	50
4.0	500	50
5.0	500	50
6.0	500	50
7.0	500	50
8.0	500	50
9.0	500	50
10.0	500	50
合計	5,000 枚	500 枚

表 7.1: データセットの枚数 (3層ニューラルネットワーク)

ためである。画像サイズが大きくなると計算機で使用するメモリ量が増えてしまう。本研究で用いた計算機の資源に対してメモリ不足とならない画像サイズとした。また、カラー画像の場合、1つのピクセルは3種類(赤、緑、青)の情報を用いて表現される。対して、グレースケールの場合は1種類でよい。つまり、画像の読込によって得られるデータ量を1/3に圧縮することができる。

学習に使用したネットワークを図 7.1 に示す。目的は機械学習を用いることができるかどうかであるため、最も簡単な構成のニューラルネットワークとした。構成が 3 層であることから、今後は 3 層ニューラルネットワークと呼ぶ。また、隠れ層のノード数を 500、100 と減らしたり 2000 と増やした 3 層ニューラルネットワークも作成した。

図 7.1 に示した 3 層ニューラルネットワークを用いて学習を行った。その結果としてトレーニングデータ (青線) と検証データ (緑線) を用いて損失関数を計算したグラフと目標値と計算値を比べて正解であった割合 (accuracy rate) のグラフを図 7.2 に示す。横軸 (epoch) はネットワークの学習回数を表し、縦軸 (loss) は損失関数を表す。損失関数は交差エントロピー関数である。右図は 2 つのデータの正解率を計算したグラフで縦軸 (acc) は正解率を表す。

結果よりトレーニングデータと検証データの損失関数の値に乖離が見られた。汎用性のあるニューラルネットワークを作成するために損失関数の値が乖離することは望ましくない。ノード数を500,100と減らして学習を行っても結果は変わらなかった。反対にノード数を増やしても結果は同様であった。これらのことから、ビームサイズの再構成を行うためには3層では足りないと結論付けた。しかし、正解率を見ると検証データで4割と出たことからビームサイズの再構成の兆しが見える。そのため、機械学習によるビームサイズ再構成は有効であると考えられる。

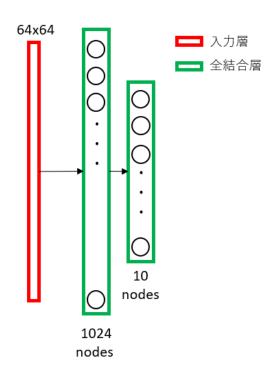


図 7.1: 3層ニューラルネットワーク

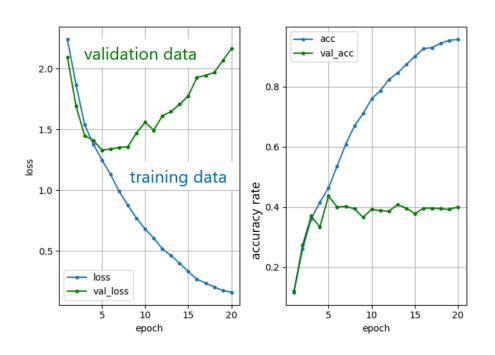


図 7.2: 損失関数と正解率のグラフ。左図、トレーニングデータ (青線) と検証データ (緑線) の損失関数を計算したグラフ。横軸 (epoch) はネットワークの学習回数を表し、縦軸 (loss) は損失関数を表す。損失関数は交差エントロピー関数である。トレーニングデータと検証データの損失関数の値に乖離が見られた。右図、2 つのデータの正解率を計算したグラフ。縦軸 (acc) は正解率を表す。

7.1.3 畳み込みニューラルネットワークを用いた検討

3層ニューラルネットワークはビームサイズの再構成ができなかったことから、より層を増やして再構成の検証を行った。0.2 倍から 3.8 倍まで 0.2 倍ずつ変化させて $(\sigma_y/\sigma_y^{TDR}=0.2,0.4,\ldots,3.8)$ 19 パターン用意した。

また、精度向上を図るため、3層ニューラルネットワークより入力画像の枚数を増やした。本研究は表 7.2 に示すように各 σ_y に対して 10,000 枚の画像を生成し、合計 190,000 枚の画像を含むデータセットを作成した。さらにこのデータセットとは別に各 σ_y に対して 5,000 枚、合計 95,000 枚のデータセットも作成した。前者をトレーニングデータと検証データとして、後者をテストデータとして使用する。トレーニングデータはデータセット全体の 9割として残りを検証データとした。

表 7.2: データセットの枚数

σ_y/σ_y^{TDR}	トレーニング+検証 データ	テストデータ
0.2	10,000 枚	5,000 枚
0.4	10,000 枚	5,000 枚
0.6	10,000 枚	5,000 枚
0.8	10,000 枚	5,000 枚
1.0	10,000 枚	5,000 枚
1.2	10,000 枚	5,000 枚
1.4	10,000 枚	5,000 枚
1.6	10,000 枚	5,000 枚
1.8	10,000 枚	5,000 枚
2.0	10,000 枚	5,000 枚
2.2	10,000 枚	5,000 枚
2.4	10,000 枚	5,000 枚
2.6	10,000 枚	5,000 枚
2.8	10,000 枚	5,000 枚
3.0	10,000 枚	5,000 枚
3.2	10,000 枚	5,000 枚
3.4	10,000 枚	5,000 枚
3.6	10,000 枚	5,000 枚
3.8	10,000 枚	5,000 枚
合計	190,000 枚	95,000 枚

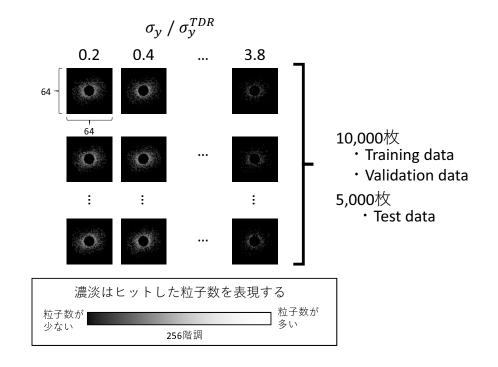


図 7.3: トレーニングデータ及びテストデータで使用した画像。画像サイズは 64×64 。ただし、上記の画像は見やすくするため拡大表示している。またヒットした粒子数をグレースケールで表現する。

学習を行うニューラルネットワークを図 7.4 に示す。例えばこのニューラルネットワークに 1 枚の画像を入力した場合図 7.5 のような過程を経る。1 枚の画像から特徴を抽出し最終的にニューラルネットワークは画像のビームサイズの確率を出力する。ニューラルネットワークは画像認識の精度の高さを競うコンペティション ImageNet[21] において画像分類部門で最も高い成績を収めた VGGnet[22] を元に、簡略化したニューラルネットワークを採用した。入力層、畳み込み層を 4 層、プーリング層を 2 層、全結合層を 2 層 (1 つは出力層) の合計 9 層のニューラルネットワークで構成した。ニューラルネットワークが複雑になると結果が出力されるまで時間がかかる。本研究の時間的都合上 VGGnet より層が浅いネットワークとした。

7.1.4 結果

図 7.6 にトレーニングデータと検証データを用いて損失関数を計算したグラフと目標値と計算値を比べて正解であった割合 (accuracy rate) のグラフを示す。

図より損失関数の値が共に 0.6 付近で収束したことから、このネットワークは未知のデータに対しても分類できることがわかった。ネットワークの設定によって学習後トレーニングデータによる損失関数の値は低くなり、検証データの関数の値が高くなる現象が現れる場合がある。この 2 つのデータで損失関数が乖離したとき、ネットワークはトレーニングデータのみ精度良く分類でき、検証データに対しては分類精度が落ちる。同様に学習後に行うテストデータに対しての分類も精度が落ちる。その結果トレーニングデータの

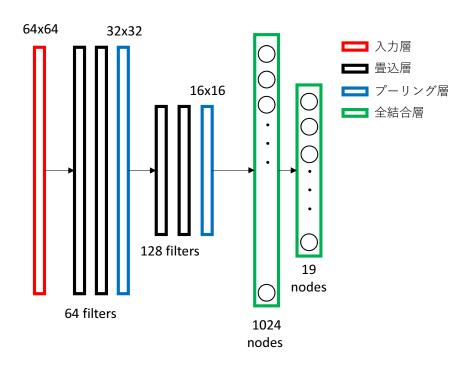


図 7.4: 畳み込みニューラルネットワーク

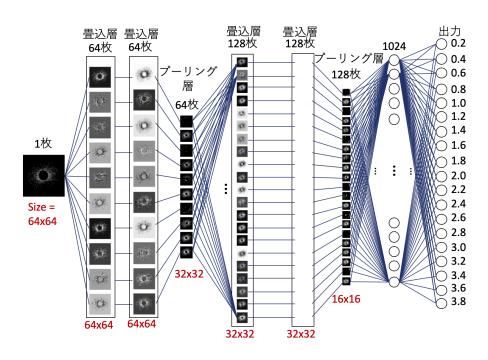


図 7.5: 1枚の画像からサイズの確率を出力するまでの過程

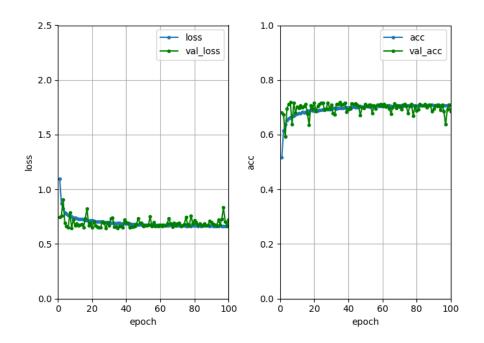


図 7.6: 損失関数と正解率のグラフ。左図、トレーニングデータ (青線) と検証データ (緑線) の損失関数を計算したグラフ。横軸 (epoch) はネットワークの学習回数を表し、縦軸 (loss) は損失関数を表す。損失関数は交差エントロピー関数である。共に 0.6 付近で収束した。右図、2 つのデータの正解率を計算したグラフ。縦軸 (acc) は正解率を表す。正解率約 7割で収束した。

分類のみに特化したネットワークになってしまうことを過学習 (over-fitting) という。本研究で採用したネットワークは過学習が起きなかったため、テストデータや実際の実験データといった未知のデータに対しても有効に動作すると考えられる。図の正解率のグラフより、2つのデータについて共に7割近い正解率を出していることからもこのことが言える。次に前述したネットワークを用いてテストデータからビームサイズの再構成を行った。このネットワークの出力は各 σ_y に対して、

$$0.2 \ \sigma_y/\sigma_y^{TDR} = n_1\%, \quad 0.4 \ \sigma_y/\sigma_y^{TDR} = n_2\%, \ \dots, \ K \ \sigma_y/\sigma_y^{TDR} = n_K\%$$

のように確率 n_K となる。このことから K σ_y^{TDR} と割合 n_K をそれぞれ掛け合わせて、全ての σ_y で加算すれば期待値を求めることができる。この期待値をビームサイズ再構成の結果とする。例として、ある画像を用いてビームサイズの確率を見たところ、図 7.7 のような結果となった。 この図から期待値 σ_y^{exp} を求めると、 $\sigma_y^{exp}=2.004$ σ_y^{TDR} となる。

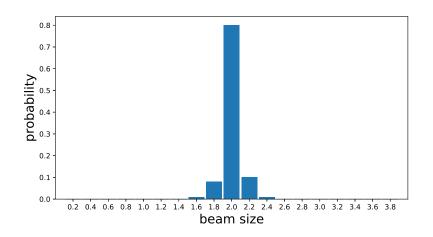
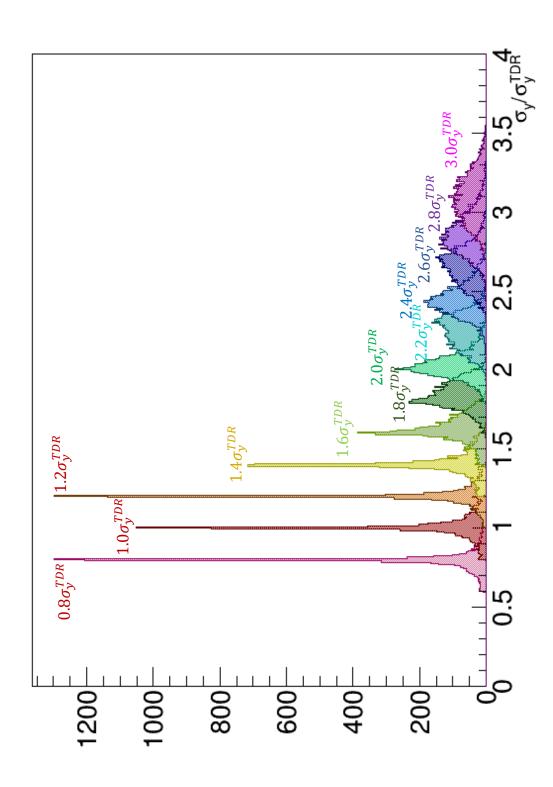


図 7.7: ビームサイズ確率分布

テストデータは $\sigma_y=0.2\sim3.8~\sigma_y^{TDR}$ までの画像が含まれる。このうち $\sigma_y=0.8\sim3.0~\sigma_y^{TDR}$ までの σ_y について期待値を計算した。その結果を図 7.8 及び表 7.3 に示す。



ビームサイズ σ_y 再構成。テストデータを元に $0.8 \le \sigma_y/\sigma_y^{TDR} \le 3.0$ のビームサイズの期待値を求めた。横軸がビームサイ 図 7.8: ビームサイズ σ_y 再構成。テスズ σ_y^{TDR} に対して何倍であるかを表す。

図 7.8 より、 σ_y が小さい値の場合目標値近傍にピークが立ち、 σ_y の増加に対して分布が広がることがわかった。これは σ_y が大きい値の時隣接する σ_y との画像の差異が少なくなり、ネットワークが目標値と異なる判断を下す割合が増えたためだと考えられる。入力に使用する画像から読み取れる項目はヒット分布の形状とヒットした粒子の数である。ただし、 σ_y を変化させたとき図 5.19 よりヒット分布の形状変化は微小である。したがって、 σ_y の再構成で重要となるのはヒットした粒子の数である。 σ_y が小さい値のときは粒子数の変化が大きいが、 σ_y が大きくなると次第に粒子の変化が少なくなる。以上のことからネットワークは目標値と異なる値を出したと考えられる。

表 7.3: 期待值

文 1.0. WIII IE		
σ_y/σ_y^{TDR}	期待値 $[\sigma_y/\sigma_y^{TDR}]$	
0.8	0.7952 ± 0.054	
1.0	1.004 ± 0.064	
1.2	1.203 ± 0.053	
1.4	1.404 ± 0.073	
1.6	1.622 ± 0.082	
1.8	1.840 ± 0.085	
2.0	2.027 ± 0.085	
2.2	2.243 ± 0.107	
2.4	2.445 ± 0.103	
2.6	2.655 ± 0.118	
2.8	2.858 ± 0.131	
3.0	3.084 ± 0.156	

7.1.5 比較

先行研究 [8] は基準となるビームパラメータの周りでテイラー展開を行い、そのテイラー展開の係数による逆行列をかけることでビームパラメータの求めた。この方法を逆行列法と定義する。本研究の再構成の精度を検討するため逆行列法によって求めた値と比較した。表 7.3 の期待値 σ_y^{exp} と逆行列法によって求めたビームサイズ σ_y' を先行研究 [8] の再構成結果のグラフから値を読み込み本結果と比較する。ただし、誤差棒はグラフから読み取ることができなかったため、最終的な再構成精度 9.9% を誤差の幅とした。ここで、先行研究は TDR ではない値 (RDR) を設定しているので σ_y と σ_y' の値は異なる。そのため、 σ_y の値で比較は行わず、 σ_y' と $\sigma_y'^{RDR}$ (RDR の設計値) から比の期待値を求めて比較を行った。図 7.9 に比較結果を示す。

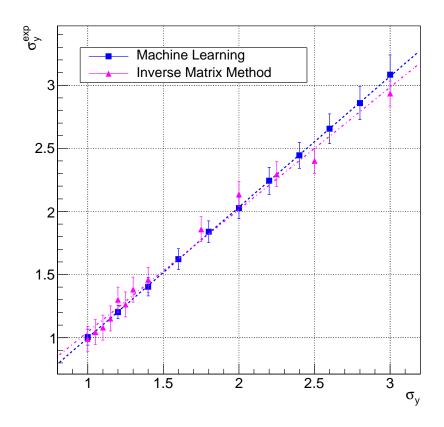


図 7.9: 逆行列法との比較。期待値を用いてフィッティングし点線で示す。

結果は逆行列法とほぼ同等な精度となった。ただし、先行研究は50バンチ分のデータを用いてビームサイズの再構成を行っているため、1バンチ分のデータを使用して再構成を行った本研究と単純に比較はできない。

7.2 ビームパラメータ (σ_x, σ_y) の再構成

次に複数のパラメータの再構成を行った。機械学習はネットワークの構成を変えれば、 複数のパラメータの測定が可能である。前節で使用したネットワークを元に、複数のパラ メータの測定が可能なように構成を変えて検証をした。

変更するビームパラメータは σ_x, σ_y として、その他のパラメータは nominal な値とする。

7.2.1 トレーニングデータ

ビームサイズ σ_r は

$$0.2 \le \sigma_x/\sigma_x^{TDR} \le 2.0$$

の範囲で $0.2\sigma_x$ ずつ変化させた。同様にビームサイズ σ_y についても

$$0.2 \le \sigma_y/\sigma_y^{TDR} \le 2.0$$

の範囲で $0.2\sigma_y$ ずつ変化させた。入力は $\sigma_x=0.2\sigma_x^{TDR}$ で固定し、 σ_y を $0.2,0.4,\ldots,2.0$ と 変化させてヒット分布を作成、次に $\sigma_x=0.4\sigma_x^{TDR}$ で固定し、 σ_y を $0.2,0.4,\ldots,2.0$ と変化させてヒット分布を作成... と σ_x が最大になるまで繰り返すことで作成した画像とする (図 7.10)。本研究は約 600,000 枚のデータセットと約 200,000 枚のデータセットを用意し、1 つをトレーニングデータ及び検証データ、もう一方をテストデータとして利用した。

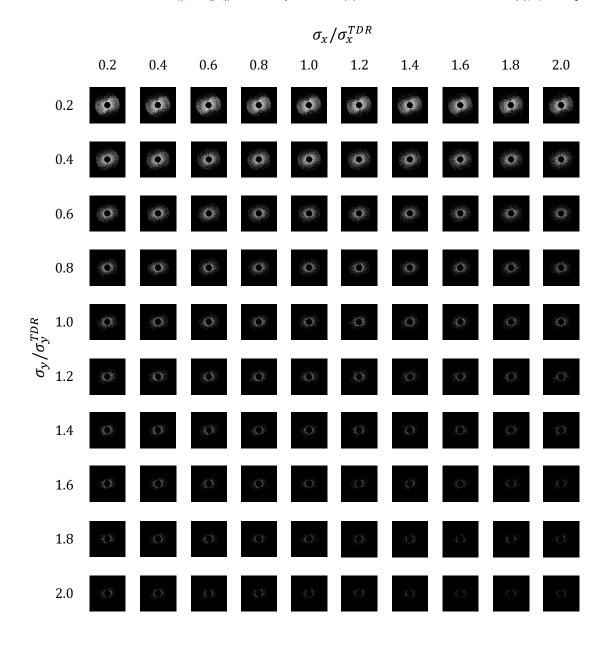


図 7.10: データセット

7.2.2 学習ネットワーク構成

学習を行うニューラルネットワークの構成を図7.11に示す。

全結合層は重みの数が畳み込み層と比べて多くなり計算に時間がかかってしまうためノードの数を100に減らした。

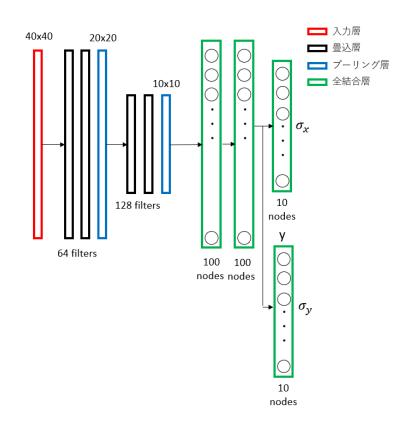


図 7.11: 使用するニューラルネットワーク構成

7.2.3 結果

図 7.12 にトレーニングデータ (青線) と検証データ (緑線) の損失関数を計算したグラフを示す。正解率のグラフはデータが取得できなかったため示していない。横軸 (epoch) はネットワークの学習回数を表し、縦軸 (loss) は損失関数を表す。損失関数は交差エントロピー関数である。

図より検証データの損失関数の値が収束しなかった。これは後述するようにビームサイズ σ_u の学習が上手くできなかったためである。

ビームサイズ σ_y のときと同様にビームサイズの再構成を行った。その結果をビームサイズ σ_x は図 7.13 に、ビームサイズ σ_y は図 7.14 に示す。

図より、 $\sigma_{x(y)}=0.2\sigma_{x(y)}^{TDR}$ 、または $\sigma_{x(y)}=2.0\sigma_{x(y)}^{TDR}$ の時は正しくビームサイズの期待値を求めることができない。学習で使用するビーム範囲は0.2が最小で2.0が最大となるため、0.2未満または2.0を超える期待値を出すことができない。例えば0.2より小さいビー

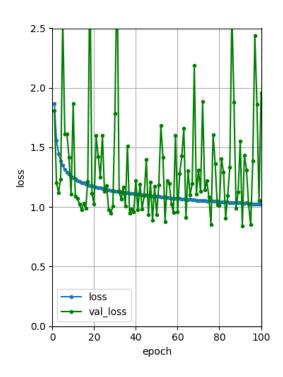
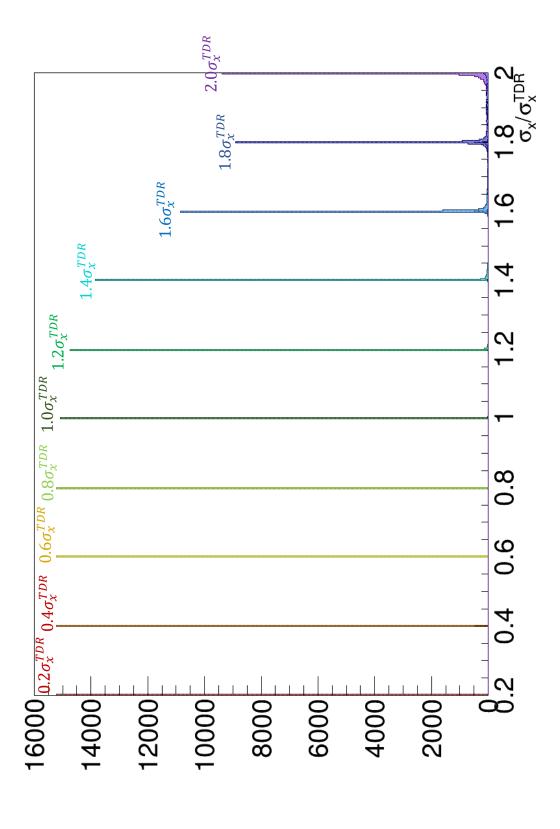


図 7.12: 損失関数のグラフ

ムサイズであってもニューラルネットワークは 0.2 が正解であると示す。同様に 2.0 以上のビームサイズであっても 2.0 と示してしまう。このことから、ビームサイズが 0.2 および 2.0 であるときの再構成は本研究のニューラルネットワークでは行うことができない。 再構成結果の図より σ_x については $0.4 \le \sigma_x/\sigma_x^{TDR} \le 1.8$ の範囲で精度良くビームサイズの再構成ができる事がわかった。 σ_x を変化させると分布図の形状が変わるため、その変化をニューラルネットワークが捉えることができたためだと考えられる。 σ_y の再構成はこれまでの結果と比べると精度が悪い。しかし、 $0.4 \le \sigma_y/\sigma_y^{TDR} \le 1.8$ の範囲でピークが $0.2\sigma_y$ ずれていることから何らかの特徴は捉えていると考えられる。ニューラルネットワークの精度を改善することができれば σ_y の再構成も可能になると予想される。



ビームサイズ σ_x 再構成。テストデータを元に $0.2 \le \sigma_x/\sigma_x^{TDR} \le 2.0$ のビームサイズの期待値を求めた。横軸がビームサイ 図 7.13: ビームサイズ σ_x 再構成。テスズ σ_{x}^{TDR} に対して何倍であるかを表す。

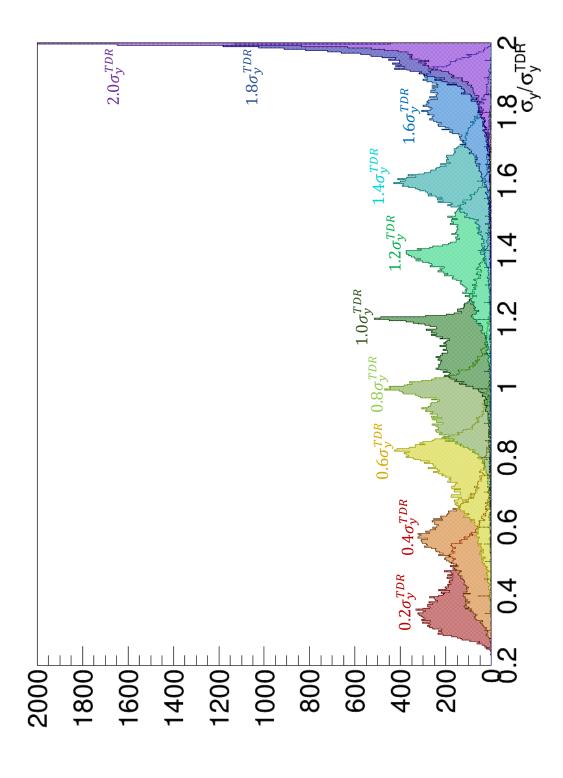


図 7.14: ビームサイズ σ_y 再構成。テストデータを元に $0.2 \le \sigma_y/\sigma_y^{TDR} \le 2.0$ のビームサイズの期待値を求めた。横軸がビームサイズ σ_y^{TDR} に対して何倍であるかを表す。

第8章 まとめ

本研究は2013年に策定されたTDRの設計値の元、ILC衝突点ビーム形状モニター(ペアモニター)の研究を行った。ビームパラメータを測定する新たな手法を提案することを目的に、本研究は機械学習を用いたビームパラメータの再構成手法を検証した。

Incoherent Pair の解析

はじめに、CAIN ソフトウェアによってビーム衝突をシミュレーションを行った。そこで生成された Incoherent Pair を ILD 内に設置されているペアモニターまで解析的に解きペアモニターのヒット分布を解析した。同電荷粒子は斥力によって反発し、異電荷粒子は引力によってビーム軸常に分布するため、ペアモニターにヒットする粒子のほとんどが同電荷粒子であることがわかった。また、先行研究と比較し解析的に解いた場合分布の広がり方が狭いことがわかった。これは anti-DID による効果であると考えられる。次に、ビームサイズ σ_x を変更させヒット分布図を解析した。ビームサイズが nominal な場合と比べて ϕ 分布の $\phi=\pm\pi/2$ 付近の割合が増えることがわかった。また、最大半径 R_{max} についても比較し σ_x によって最大半径の取り得る値が変化することがわかった。ビームサイズ σ_y を変更させヒット分布図を解析した。ヒット分布については $\phi=\pm\pi/2$ 付近で粒子の割合の減少が見られた。しかし、 R_{max} は差が微小であった。以上より、 σ_x および σ_y は値を変化させるとヒット分布の形状が変化するため機械学習に入力として利用することが可能であることがわかった。

機械学習を用いたビームパラメータ再構成の検討

次に、機械学習でビームサイズ σ_y の再構成を行った。機械学習のネットワークは畳み込みニューラルネットワークを採用した。これは、入力層、畳み込み層、プーリング層、出力層からなる合計 9 層のニューラルネットワークで VGGnet を元に作成した。ネットワークは Tensorflow および Keras を用いてこのプログラムを作成する。入力としてペアモニターのヒット分布を 190,000 枚作成し、トレーニングおよび検証に利用した。その結果を図 8.1 に示す。

本研究の手法を用いてのビーム再構成ができることがわかった。

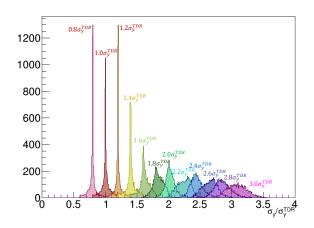


図 8.1: ビームサイズ σ_y 再構成。

また、ビームサイズを σ_x , σ_y 同時に再構成できるか検証した。結果を図 8.2 に示す。結果は σ_x については精度良く再構成を行うことができた。しかし σ_y については精度がこれまでと比べて低いため、ニューラルネットワークの改善が必要である。以上の結果から、ニューラルネットワークの調整を行えば、機械学習を用いたビームサイズ再構成は可能である。

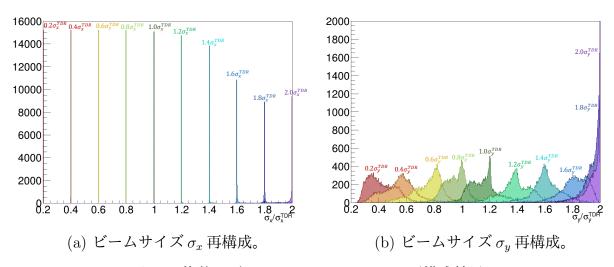


図 8.2: 複数のビームパラメータ (σ_x, σ_y) 再構成結果

謝辞

本研究を進めるにあたって、多くの皆様にお世話になりました。はじめに佐貫先生には大変お世話になりました。研究でわからないことがあると、助言をいただくこともあれば、その内容について詳しく知っている先生を紹介してもらったりするなど、適切なご指導をしていただき、有意義な研究をすることができました。また、LCWSで発表を行うことが決まったとき、発表用スライド作成が遅れたにもかかわらずぎりぎりまでスライドの修正を手伝っていただきました。本当にありがとうございました。加えて、山本先生を始めとする素粒子実験研究室の先生方にもお世話になりました。山本先生には素粒子物理の基礎から懇切丁寧に教えていただき大変感謝しています。また、与那嶺先生にはROOTの使い方やCAINを使いやすくするスクリプトを作成して頂きました。そしてセミナーでは様々な助言を頂くことでき、物理に対する理解が深まりました。長嶺先生にはネットワークやホームページサーバについて、石川先生にはFPGA研修でお世話になりました。重ねて感謝申し上げます。

岩崎先生を代表とする阪大 DNN グループの皆様にもお世話になりました。機械学習について左も右もわからないなかで、研究を進めるために必要な基本的なことから教えていただきました。最後に素粒子実験研究室の先輩、同期、後輩そして事務の皆様、本当にありがとうございました。

参考文献

- [1] T. Behnke *et al.*, "The International Linear Collider Technical Design Report Volume 1: Executive Summary," arXiv:1306.6327 [physics.acc-ph].
- [2] T. Behnke *et al.*, "The International Linear Collider Technical Design Report Volume 4: Detectors," arXiv:1306.6329 [physics.ins-det].
- [3] M. A. Thomson, "Particle Flow Calorimetry and the PandoraPFA Algorithm," arXiv:0907.3577 [physics.ins-det].
- [4] 木村 嘉孝、『第7巻 高エネルギー加速器』, 共栄出版株式会社, 2008
- [5] T. Tauchi and K. Yokoya, "Nanometer-beam-size measurement during collisions at linear colliders," Phys. Rev. E 51, 6119 (1995).
- [6] CAIN web site. https://ilc.kek.jp/~yokoya/CAIN/Cain242/
- [7] K.Yokoya, "A computer simulation code for the beam-beam interaction in linear colliders," KEK Report 85-9 October (1985).
- [8] 佐藤 優太郎, "国際リニアコライダーのための衝突点ビーム形状モニターの研究開発," 修士論文, 東北大学, (2009)
- [9] J. Brau *et al.*, "ILC Reference Design Report Volume 1 Executive Summary," arXiv:0712.1950 [physics.acc-ph].
- [10] 山口 信二郎, "国際リニアコライダーのビーム形状測定のためのペアモニターの研究"修士論文, 東北大学, (2015)
- [11] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers," in IBM Journal of Research and Development, vol. 3, no. 3, pp. 210-229, July 1959. doi: 10.1147/rd.33.0210
- [12] Rosenblatt, F., "The perceptron: A probabilistic model for information storage and organization in the brain," Psychological Review, **65**(6), 386-408 (1958).
- [13] Vapnik, V. and A. Lerner, "Pattern recognition using generalized portrait method," Automation and Remote Control. 24, 774-780, (1963).

[14] Fukushima K., "Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," Biol Cybern. **36**(4),193-202 (1980).

- [15] Alex Krizhevsky and Sutskever, Ilya and Hinton, Geoffrey E, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems 25, 1097–1105, (2012).
- [16] Quoc V. Le *et al.*, "Building high-level features using large scale unsupervised learning," arXiv:1112.6209 [cs.LG].
- [17] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, Ruslan Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," Journal of Machine Learning Research 15, 1929–1958, (2014).
- [18] S. Ioffe, et al., "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv:1502.03167 [cs.LG].
- [19] Xiang Li, Shuo Chen, Xiaolin Hu, Jian Yang, "Understanding the Disharmony between Dropout and Batch Normalization by Variance Shift," arXiv:1801.05134 [cs.LG].
- [20] Tensorflow web site. https://www.tensorflow.org/
- [21] ImageNet web site. http://www.image-net.org/challenges/LSVRC/2014/
- [22] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556 [cs.CV].