

CNNにおける認識率向上のための 層数とloss関数の選定

電気電子情報工学科

木下研究室

遠藤 颯汰(201604424)

～研究背景～

深層学習を利用した画像認識の一分野である
「クラス分類」の認識率の向上のために
様々なアプローチが提案されている

～先行研究～

Kaggleの競技にてDeep Seaが採用した、以下の方法を利用

Self Training
Stacked Generalization
複数モデルの利用

- ・3つの方法を組み合わせてモデル平均の正解率が4.8ポイント上昇
- ・しかし、データセット枚数が6倍となる事に比例し、実行時間も8～9倍(約9時間)となってしまうあまり効率的とは言えない
- ・本研究では実行時間に一定の制約を設けた上で、認識率の向上を図る

～目的～

以下の2通りの方法を提案

①畳み込み層を追加し、正解率を比較、効率的な認識率の向上が可能な層数の選定

②確率的勾配効果法の学習係数 ε に着目、9層CNNで層数を増やすことなく認識率が向上する最適値を選定

・実行時間、容量の観点より本研究では9層CNNをベースモデルに使用する

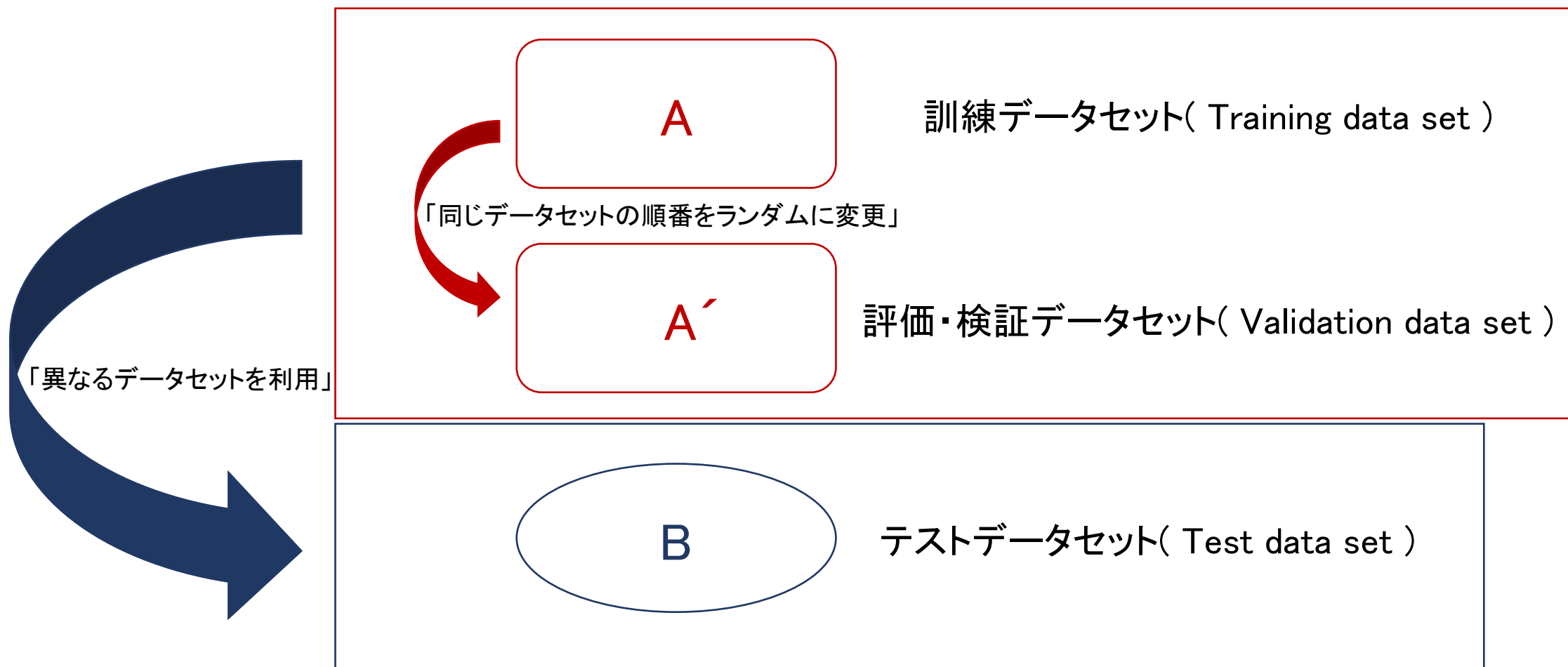
～使用する9層CNNの概要～

9層

- 畳み込み層(conv) 6層
- 全結合層(fc) 3層
- 9層のCNN を利用
- プーリングサイズ 2×2 (MaxPooling)
- 224×224 ピクセル画像3枚(RGB)入力
- 教師あり学習
- データセット「Caltech101」6クラス

入力層		(層番号)
3 × 3 conv,32		1
3 × 3 conv,32		2
maxpool,/2		
3 × 3 conv,64		3
3 × 3 conv,64		4
maxpool,/2		
3 × 3 conv,128		5
3 × 3 conv,128		6
maxpool,/2		
fc 1024		7
fc 1024		8
fc 6		9
出力データ		

～Data set の説明～



① 畳み込み層追加・実験

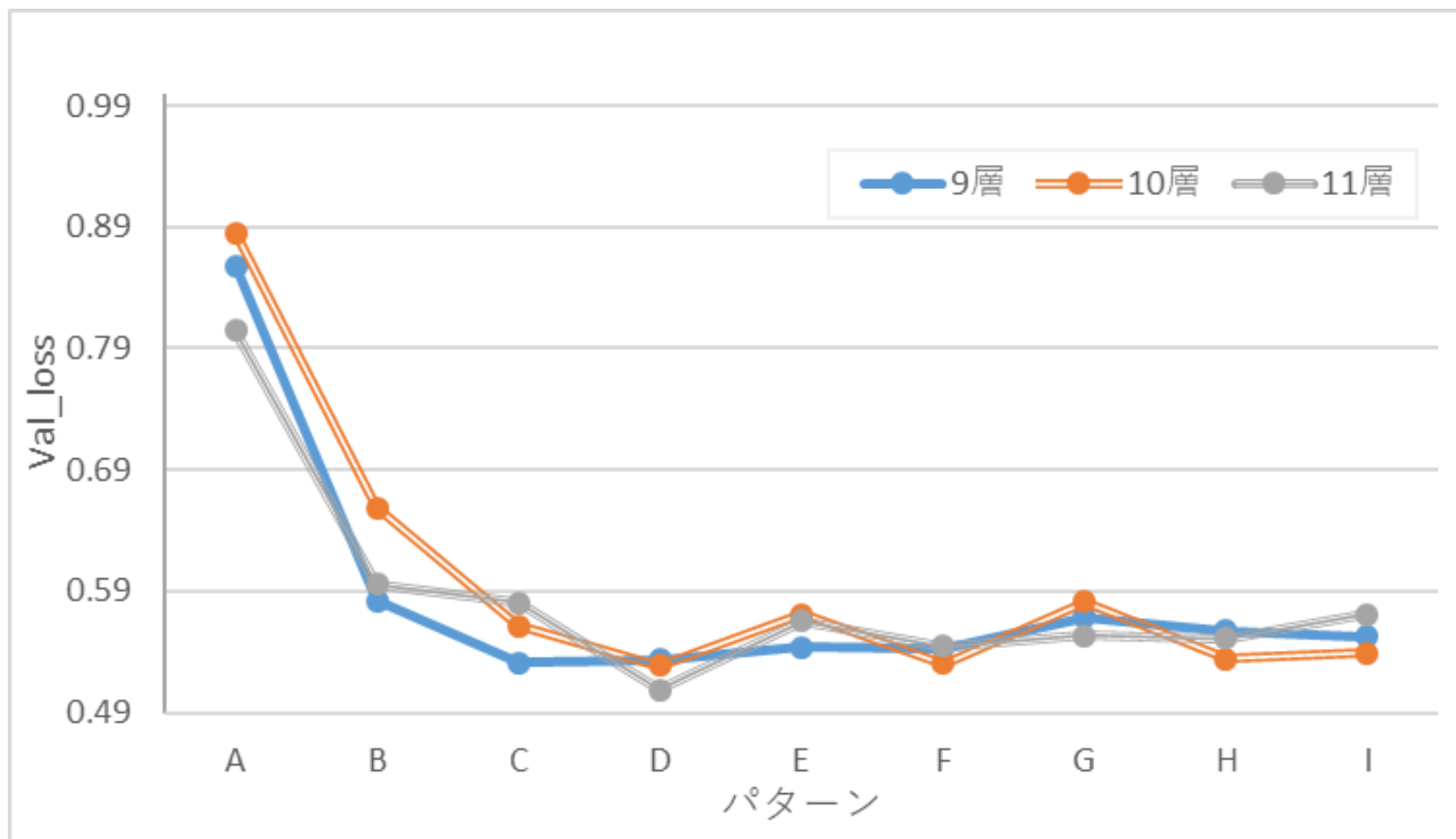
・ (フィルターサイズ 3×3)(256枚) 畳み込み層を1層ずつ9層CNNに追加

・ プーリング層1層追加

・ 各層数にて9回のテスト

10層		11層	
入力層		入力層	
	(層番号)		(層番号)
3 × 3 conv,32	1	3 × 3 conv,32	1
3 × 3 conv,32	2	3 × 3 conv,32	2
maxpool,/2		maxpool,/2	
3 × 3 conv,64	3	3 × 3 conv,64	3
3 × 3 conv,64	4	3 × 3 conv,64	4
maxpool,/2		maxpool,/2	
3 × 3 conv,128	5	3 × 3 conv,128	5
3 × 3 conv,128	6	3 × 3 conv,128	6
maxpool,/2		maxpool,/2	
3 × 3 conv,256	7	3 × 3 conv,256	7
maxpool,/2		3 × 3 conv,256	8
fc 1024	8	maxpool,/2	
fc 1024	9	fc 1024	9
fc 6	10	fc 1024	10
		fc 6	11
出力データ		出力データ	

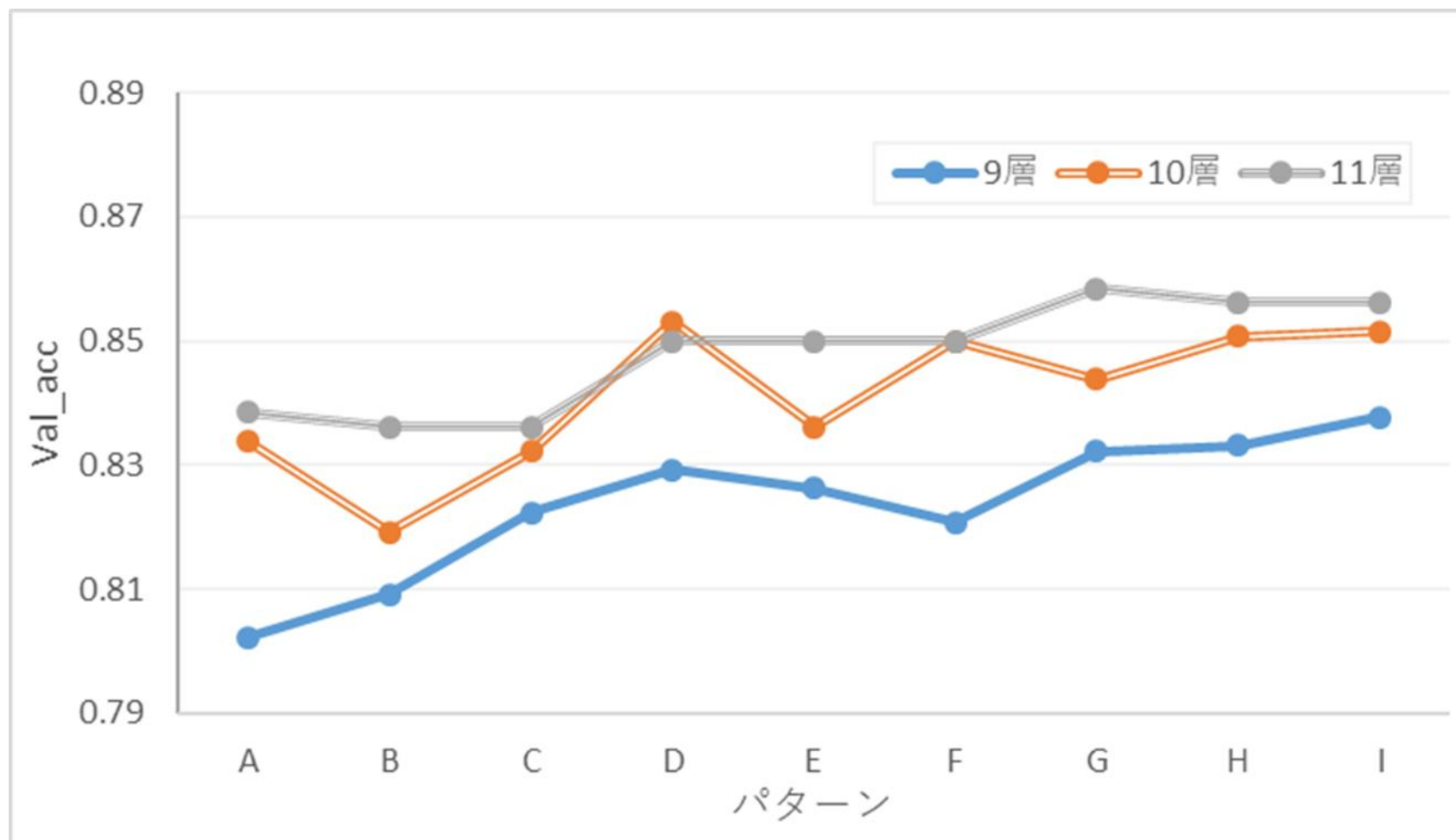
①各実験パターンにおける9.10.11層の Val_lossの比較(Validation data set 評価結果)



←低層パラメータ変更

高層パラメータ変更→

①各実験パターンにおける9.10.11層の Val_accの比較(Validation data set 評価結果)



←低層パラメータ変更

高層パラメータ変更→

① 9パターンの実験におけるVal_loss/Val_acc (Validation data set 評価結果)

	最良値		平均値	
	val_loss	val_acc	val_loss	val_acc
9層	0.5311	0.8377	0.5856	0.8237
10層	0.5285	0.8531	0.5989	0.8412
11層	0.5088	0.8585	0.5861	0.8480

※ Val_accはより大きい値が良く、Val_lossはより小さい値が良い

① 9.10.11層のAccuracy (Test data set 結果)

		テストデータセットの真のクラス						合計	
		airplanes	motorbikes	faces_easy	watch	leopards	bonsai		
9層	推測した クラス	正解	383	587	324	49	137	20	1500
		不正解	257	51	24	142	23	82	579
		計	640	638	348	191	160	102	2079
Accuracy		60%	92%	93%	26%	86%	20%	72%	

		テストデータセットの真のクラス						合計	
		airplanes	motorbikes	faces_easy	watch	leopards	bonsai		
10層	推測した クラス	正解	389	587	323	45	132	18	1494
		不正解	251	51	25	146	28	84	585
		計	640	638	348	191	160	102	2079
Accuracy		61%	92%	93%	24%	83%	18%	72%	

		テストデータセットの真のクラス						合計	
		airplanes	motorbikes	faces_easy	watch	leopards	bonsai		
11層	推測した クラス	正解	362	596	308	34	129	22	1451
		不正解	278	42	40	157	31	80	628
		計	640	638	348	191	160	102	2079
Accuracy		57%	93%	89%	18%	81%	22%	70%	

① 実験結果(畳み込み層追加)

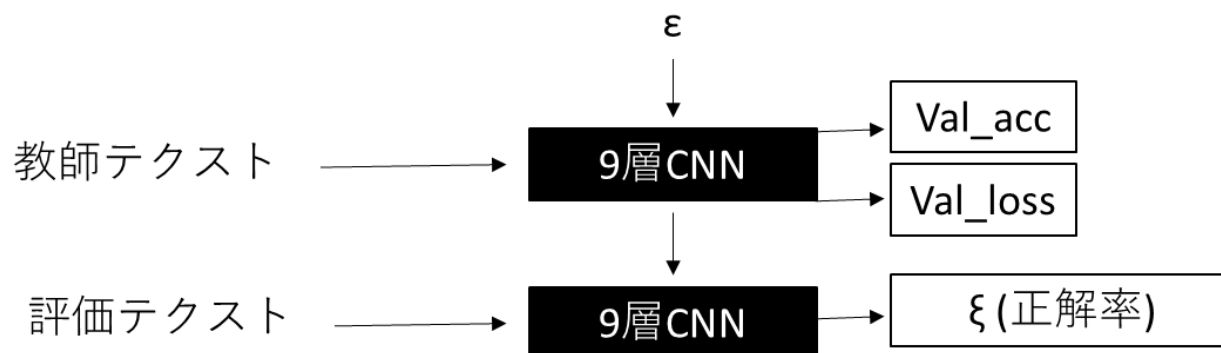
Val_loss/Val_acc(Validation data set 評価結果)

ほとんどのパラメータの組み合わせにおいて値の向上が見られた

Accuracy(Test data set 結果)

10層11層は9層に比べ優位性が見られなかった

② 学習係数 ε の最適値選定モデル

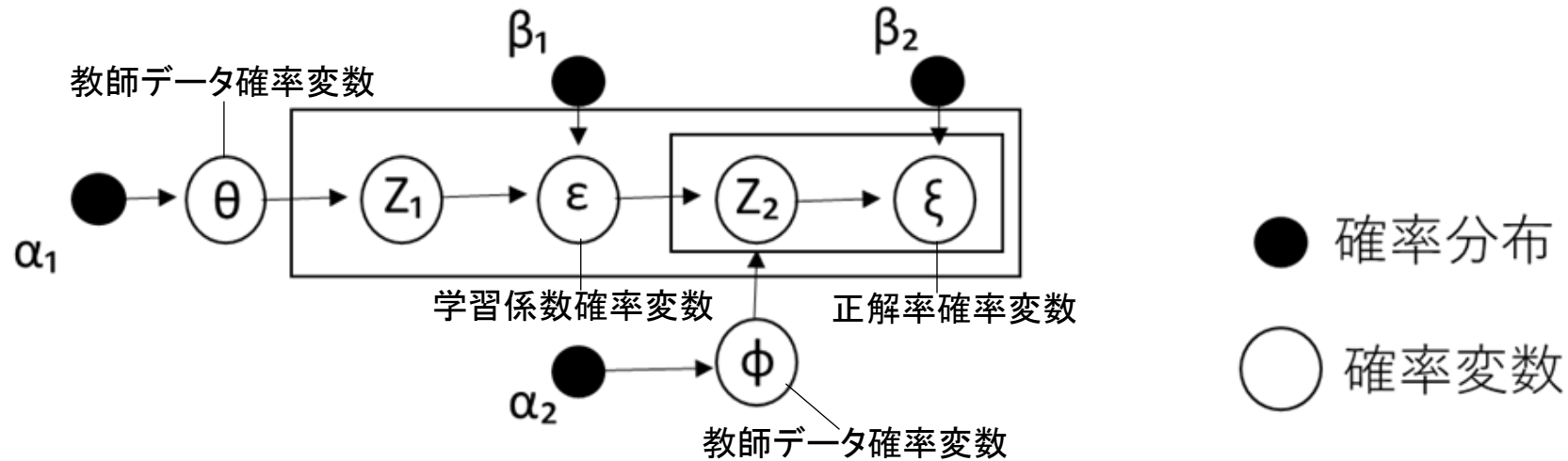


観測可能なデータ

(教師テキスト/評価テキスト/ ε の値/正解率 ξ)

- ・ 最適な正解率 ξ となる ε の値を求めるには(教師テキスト/評価テキスト)も複雑に関わってくる

② 学習係数 ε のベイズ確率モデル



・上図は、 ε の最適値選定モデルにおいて、 Z_1 と Z_2 をそれぞれ LDA(Latent Dirichlet Allocation)の潜在的確率変数とみなし、2つの LDAを連結するモデルによってベイズ推定するモデル

・本研究では、上記ベイズ推定の枠組みの中において、学習係数 ε が正解率 ξ に及ぼす効果を確認する目的で、手動による事前準備実験として、8通りの値を1つ1つ当てはめることで、 ε の最適値の選定を行う

② 学習係数 ε 選定・実験

- ・機械学習を実施する前にブロック図において、学習係数 ε が正解率 ξ に及ぼす効果を確認する目的

～手動による事前準備実験～

0.1～0.000000001 まで8通りの値を1つ1つ当てはめることで

ε の最適値の選定

- ・各値による、異なる条件5パターンでの実験を行い、正解率の平均を集計、比較

② 5パターンでのAccuracy(Test data set 結果)

	airplanes	motorbikes	Faces easy	watch	leopards	bonsai	Accuracy	
$\epsilon=0.001$	A	60	92	93	26	86	20	72%
	B	57	93	89	18	81	22	70%
	C	61	92	93	24	83	18	72%
	D	59	92	88	9	74	1	68%
	E	60	92	93	24	86	17	72%
	airplanes	motorbikes	Faces easy	watch	leopards	bonsai	Accuracy	
$\epsilon=0.01$	A	63	89	90	47	88	50	75%
	B	66	92	94	60	91	66	80%
	C	60	92	97	46	89	61	77%
	D	63	89	95	51	91	53	77%
	E	65	90	94	66	88	63	79%

② 実験結果(ε の最適値選定)

- ・0.001と0.01以外の6通りを当てはめた場合、エラーや50%未満の結果となった
- ・9層のCNNにおいては、 ε の値に0.01を使用する事で正解率が平均6.8ポイント向上
- ・先行研究に比べて1/8の実行時間で2ポイント上回る結果となった

9層CNNで最適な ε の値と考えられるのは、 $\varepsilon = 0.01$ である

～今後の課題～

- ・本研究での ε の値の選定は9層のCNNに限定された最適値だったため、この結果を層数の多い学習済みモデルに応用し、同じように認識率の向上が見られるのか改めて検討する
- ・使用した6クラスの画像データセット以外の画像データを用いて実験を行うことで同じような検証結果が見られるのか、実験を行う

～結論～

- ・本研究では画像認識率向上を目的として、 ε の最適値選定を提案することで**効率的な認識率の向上**が可能となった
- ・今後の課題はいくつか見つかったが、**本研究の目的はおおむね達成できた**と言える
- ・本研究のような画像認識の技術向上や効率化を目的とした研究は、今後の産業革命の分野においてより重要になってくると考えられる