

The logo for NTT DATA, featuring the company name in a bold, white, sans-serif font against a blue background.

NTT DATA Mathematical Systems Inc.

A decorative graphic on the left side of the slide, consisting of a grid of squares in various shades of blue, with some squares containing semi-circular shapes in orange, yellow, and teal.

AutoDL 納品レポートサンプル

株式会社NTTデータ数理システム

目次

- 概要
- 納品ファイルリスト
- 実験設定
- 実験結果
 - 学習曲線
 - 的中率
 - Confusion Matrix
 - 誤判別画像
- 考察

- 受領データの分割
 - 受領データを8:1:1に分割し、学習用、検証用、評価用とします。
 - 学習用、検証用、評価用で各ラベルの割合は同じになるように分割しています。
- 処理の流れ
 1. 学習用画像を用いてモデルを学習
 2. 検証用画像に対する精度が収束した段階で学習を打ち切り
 3. 学習したモデルを使用して、評価用画像に対する精度を本資料にて報告

納品物リスト

- 実験レポート.pptx
 - 本資料
 - 実験設定、モデル精度に関するレポート
- data_split.csv
 - 学習用、検証用、評価用に用いた画像ファイル名のリスト
 - 1列目:画像ファイル名
 - 2列目:データ区分(0:学習用、1:検証用、2:評価用)
- 誤判別画像フォルダ
 - 評価用画像の内、モデルが予測を間違えた画像群
 - 正解ラベル名前でフォルダ分けされています。
 - 画像ファイル名の先頭が予測ラベル名です。
- prediction.csv
 - 検証用データに対する、各ラベルの予測確率リスト
- deployフォルダ
 - 学習済みモデルおよびモデル実行用のコード群

項目	内容	備考
受領データ内容	犬、猫の写真群	
総画像数	10000	ラベル毎の画像数は次スライド参照
実験環境	aws: p2.xlarge	Amazon Web Services https://aws.amazon.com/jp/p2instances p2インスタンス https://aws.amazon.com/jp/ec2/instance-types/p2/
モデル	ResNet18	https://arxiv.org/pdf/1512.03385.pdf

- ラベル毎の画像数

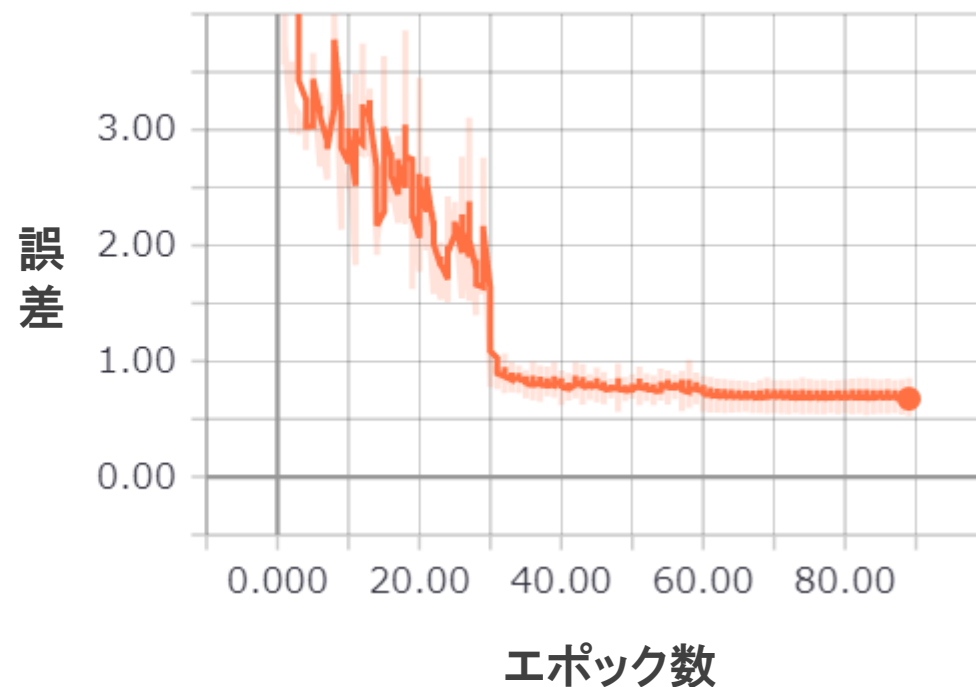
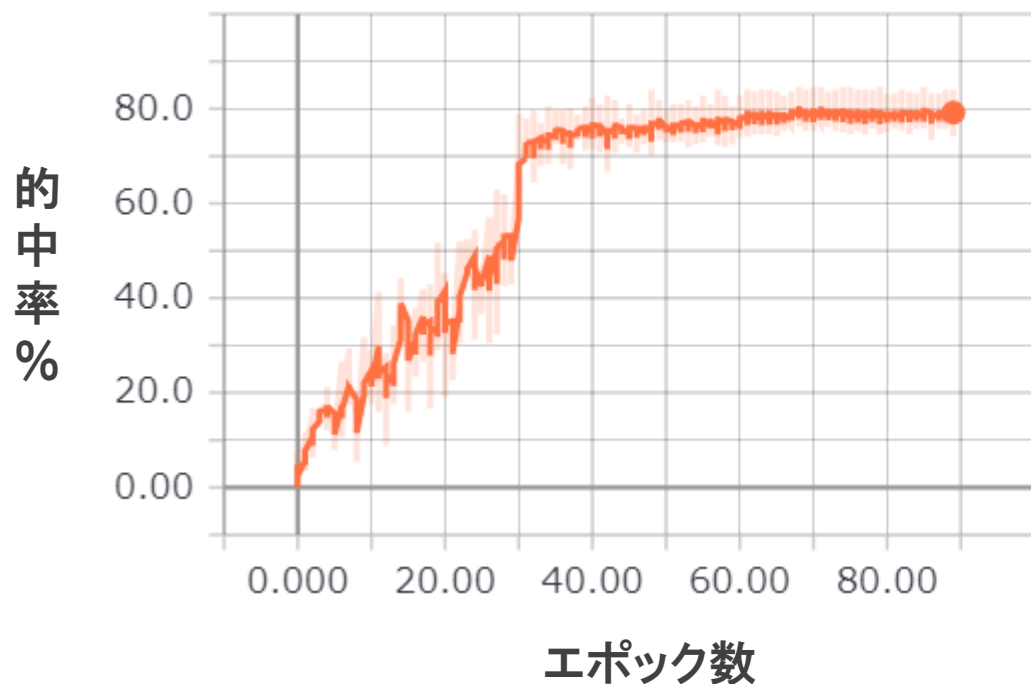
ラベル名	枚数
シャム	1000
ロシアンブルー	1000
柴犬	1000
パグ	1000
その他	6000

- 学習用、検証用、評価用データ

- 学習用、検証用、評価用に用いた画像ファイル名に関しては、“data_split.csv”を参照してください。

- 学習曲線

- 検証用データに対する精度のグラフです。
- おおむね80エポック(80回ずつ画像を読み込んだ状態)で学習が収束しているのが確認できます。



- 的中率

項目	値	備考
overall accuracy	80.00%	予測の的中率
average accuracy	75.00%	正解ラベル毎の正解率の平均
top-3 accuracy	98.32%	確率の高い上位3つのラベルを予測した場合の的中率

- Confusion Matrix

- 各ラベル毎に予測された画像の枚数を表すマトリクスです。

		予測ラベル					正解ラベル毎の 正解率
		シャム	ロシアン ブルー	柴犬	パグ	その他	
正 解 ラ ベ ル	シャム	50	1	2	0	47	50.00%
	ロシアンブルー	0	80	0	0	20	80.00%
	柴犬	0	0	80	8	12	80.00%
	パグ	0	0	7	80	13	80.00%
	その他	60	10	10	10	510	85.00%
予測ラベル毎の 正解率		45.45%	87.91%	80.81%	81.63%	84.72%	

- 誤判別画像群
 - 誤判別画像群は”誤判別画像”フォルダを参照してください。
 - 正解ラベル名毎にフォルダ分けされています。
 - 各画像ファイル名が“予測ラベル名_元の画像ファイル名”となっています。
- 予測確率リスト(prediction.csv)
 - 評価用データに対する、各ラベルの予測確率です。
 - 1列目が画像ファイル名、2列目以降が各ラベルに対応します。