

FUJITSU AI ソリューション

Zinrai ディープラーニング システム

制限事項・注意事項

Copyright 2018 FUJITSU LIMITED

P3KD-1142-03

本書では、FUJITSU AI ソリューション Zinrai ディープラーニング システムにおける、制限事項・注意事項について説明します。各機能の詳細は、『機能説明書』を参照してください。

2018年10月第3版

制限事項

バッチ型学習

- バッチ型学習実行時に指定できる GPU 数は 1、2、4 です。8 は指定できません。
- 学習ウィザードの各入力画面で操作しているときに、別ブラウザでログイン画面を立ち上げると、元の画面では「次へ」ボタンを押下したときに画面上に「undifined」と表示され、あとから立ち上げたログイン画面からログインを試みると「403 Access Denied」が表示され両画面ともアクセスできなくなります。

【対処方法】

アクセスできなくなった場合は、再ログインしてください。

新規学習、追加学習、NN 最適化

- 新規学習・追加学習・NN 最適化で GPU を 2 つ以上使用した場合、GPU1 つで学習を行った場合よりも早く学習が収束しますが、GPU1 つに相当する学習を終えたのちも (accuracy 値 (正解率) がモデル相応の水準に達したあとも) 学習を継続します。

複数 GPU を使用した場合でも、最終的に GPU1 つで学習を行った時間と同程度経過したのち学習が終了します。なお、GPU 数に応じてリニアに近い学習の高速化が行えるか (GPU 数に応じて学習の収束が早まるか) はモデルのネットワークに依存します。もともと、GPU 数に応じて学習が高速化するようにチューニングされていないネットワークの場合は、使用した GPU 数に応じて学習の収束、および 1GPU 相当の学習を終えたのち学習を継続し学習が終了するまでの時間が長くなる場合があります。

【回避方法】

新規学習・追加学習の場合、グローバルナビゲーションで [管理] - [モデル] を選択し、モデルの詳細情報より、グラフの accuracy 値 (正解率)、テストロスの収束状況を定期的を確認してください。期待する水準の accuracy 値に達している場合は、グローバルナビゲーションで [管理] - [モデル] より、学習の中断を行ってください。

NN 最適化に関しては、自動生成したハイパーパラメーターに対する学習を個別に中断できないため、回避方法はありません。

- 新規学習・追加学習のスナップショット、および NN 最適化の実行結果のファイルサイズが大きい場合、ダウンロードが失敗することがあります。ダウンロードファイルの準備処理が 1 分以内に完了しない場合、ダウンロード処理が開始前に打ち切れ、「An error occurred.」または「502 Bad Gateway」のエラーがブラウザ画面に表示されます。

【回避方法】

上記ファイルのサイズを小さくするため、以下の方法で学習を行ってください。

- AlexNet や LeNet の代わりに GoogleNet を使用してください。
- Crop Size は 223 以上、かつ 256 以下で設定してください。

NN 最適化

- NN 最適化のワーキングセットに対し追加学習を行うことはできませんが、ワーキングセットの一覧では、NN 最適化のモデルに対して「追加学習」ボタンが有効になっています。このとき「追加学習」を実行しても、学習ウィザードのスナップショットの選択画面で、スナップショットが表示されません。

【回避方法】

回避方法はありません。NN 最適化のワーキングセットに対し追加学習の操作をしないでください。

教師データ

- 教師データの labels.txt に 4,500 以上のラベルを含む場合、学習後の認識（1 枚指定、複数枚指定）に失敗し、「予期しないエラーが発生しました」と表示されることがあります。

【回避方法】

教師データの labels.txt に含めるラベル数は 4,500 未満にしてください。

データの管理

- 管理メニューで管理される、「ワーキングセット」、「モデル」、「データセット」、「認識結果」の各データ一覧画面より、一括で数十件のデータを同時に削除すると、エラー画面が表示され、正常に削除完了しないことがあります。

【回避方法】

各データの一覧での一括削除数は 10 以下としてください。

知識ライブラリ

- 知識ライブラリの所持データ情報編集画面で、データ種別を「モデルセット」または「その他」から「教師データ」または「Edge 用データ」に変更した場合、学習ウィザードのデータセットの作成画面に表示される画像ファイル数の値が、画像ファイル数ではなく、全ファイル数で表示されます。

<画像ファイル数、全ファイル数について>

全ファイル数：すべてのファイルをカウントした数

画像ファイル数：拡張子が jpg、jpeg、png のファイルをカウントした数

<教師データのディレクトリ構成の例>

labels.txt

train/airplane/001.jpg

 /002.jpg

car/001.jpg

 /002.jpg

上記の例の場合、全ファイル数は 5、画像ファイル数は 4 です。

【回避方法】

教師データをアップロードする場合、データ種別を「教師データ」または「Edge 用データ」に指定してください。

データ種別に「モデルセット」または「その他」を選択してアップロードした場合、一度データを削除したあと、データ種別を「教師データ」または「Edge 用データ」に指定して再度アップロードを行ってください。

エッジ連携

- エッジモデルを Android、iOS へダウンロードする場合、ダウンロードに失敗することがあります。エッジモデルのサイズが大きい場合（100MB 以上が目安のサイズになります）、準備処理に時間がかかります。準備処理が 10 秒以内に完了しない場合、ダウンロード処理が開始前に打ち切れ、ダウンロードに失敗します。

【回避方法】

エッジモデルのサイズを小さくするため、以下の方法で学習を行ってください。

- AlexNet や LeNet の代わりに GoogleNet を使用してください。
- Crop Size は 223 以上、かつ 256 以下で設定してください。

- バッチ型学習にて、LeNet のように Data Layer の transform_param の scale 値に 1 以外（指定しない場合はデフォルト値 1 となります）を指定して学習したモデルをエッジで利用すると、正しく推論できず誤った結果となります。

【回避方法】

エッジ推論で利用する学習モデルは、GoogleNet, AlexNet を使用してください。

注意事項

学習ウィザード

■ 学習ウィザード内は先に進むだけで、戻る機能はありません。

- 各モーダルでのキャンセル時およびエラーメッセージ表示時は、ウィザードが終了しワーキングセット一覧に戻ります。
- ウィザード内のボタン実行で保存されるデータについて
 - 各管理画面でだけ削除可能
 - 保存ボタン一覧（○：保存、－：対象外）

画面	ワーキングセット作成	ネットワークコード編集	ネットワーク選択	データセット作成	学習設定	学習設定
ボタン名	次へ	確定	次へ	確定	ジョブランク選択	実行
ワーキングセット	○	－	－	－	－	－
モデル（新規）	－	○	－	－	○	○
モデル（流用）	－	－	○	－	○	○
データセット	－	－	－	○	－	－

■ 学習ウィザードでデータセット作成し学習を行う場合に、データセット作成が完了していないときはデータセット選択画面の画像数が0枚と表示されます。

【対処方法】

データセット作成完了後は実際に使用した画像数が表示されます。作成完了後に確認してください。

■ コード編集で「確定」ボタンを押下した場合、コード編集エラーがあっても必ずネットワーク（モデル）データが作成されます。
エラーのあったネットワークデータを参照しても、コードが空で表示されます。

【対処方法】

コードエラーがあるデータは使用できませんので、削除する場合はグローバルメニューの[管理] → [モデル] の一覧から削除してください。

■ 学習ウィザードのデータセット選択画面において、「ストップ」または「エラー」のデータセットを選択すると、学習のサブミットジョブが「準備中」の状態のままとなり、学習ジョブは実行されません。

【対処方法】

学習を実行するデータセットとして、ステータスが「ストップ」および「エラー」のデータセットを選択しないでください。

- ネットワークの選択画面の選択情報上に表示される「メモ」は、「学習ウィザード 学習設定」画面で入力されるメモを表示します。
従って、ネットワークの選択画面で学習済みネットワークを選択した時は、学習時に設定したメモが表示されます。
しかし、新規作成後にネットワーク選択画面から新規ネットワークを選択した場合、メモ入力フェーズが無いため、空白表示となります。

【回避方法】

回避方法はありません。

- 学習回数（Training epochs）の設定が大きい場合、次の操作で表示速度が極端に遅くなる場合があります。

- 管理→モデル一覧→モデル詳細表示時
- 認識→モデル一覧→認識設定表示時
- 管理→認識結果一覧→詳細表示時
- 追加学習→スナップショットの選択表示時
- NN 最適化→学習設定表示時

【回避方法】

Solver 設定 Training epochs（学習回数）は 250 以下になるよう設定してください。

ダイアログ表示

- 操作画面のボタン押下により呼び出されるモーダルダイアログを ESC キー操作で閉じた場合、元のダイアログ表示させたボタンが押下できない状態のままとなってしまいます。

【回避方法】

モーダルダイアログを ESC キー操作で閉じないでください。

ESC キーでモーダルダイアログを閉じてボタンが押せない状態になった場合は、ページの再読み込みを行ってください。

学習結果の確認（画像認識）

- 学習結果の確認で使用するモデルの Accuracy 値が低い（10% 程度）場合、画像認識がエラー終了する場合があります。
Accuracy 値が高い #epoch を選択してください。また、Accuracy 値は次の手順で確認してください。

【手順】

- 1 [管理] メニューより、[モデル] の項目を選択し、モデル一覧画面を開く。
- 2 モデル一覧画面より、学習結果の確認を行うモデルのリンクをクリックし、モデル詳細（新規学習・追加学習）画面を開く。
- 3 スナップショット情報を参照して、各スナップショット名の Accuracy 値を確認する。

学習結果の確認

- スナップショットの Accuracy 表示において、同一スナップショットであるのに画面によって値に誤差が生じます。
小数点以下の端数処理方法に違いがあるためです。

各画面における端数処理は以下のようになります。

管理 モデル詳細 ⇒ 四捨五入

管理 認識結果 ⇒ 切捨て

学習ウィザード（追加学習）ネットワーク選択 ⇒ 切捨て

認識 認識設定 ⇒ 切捨て

追加学習

- 追加学習時に学習の設定で Crop Size を学習済みモデルと異なる値にすると、学習エラーになります。
学習エラー時は、Caffe ログに "shape mismatch" のエラーが記録されます。

エラーの例：

```
Fmdd hh:mm:ss.nnnnnn 31 net.cpp:829] Cannot copy param 0 weights from layer 'fc6'; shape mismatch. ~
```

(注) Fmdd hh:mm:ss.nnnnnn

Fmdd の mm は月、dd は日、hh:mm:ss.nnnnnn は時刻を表します。

【回避方法】

学習の設定で Crop Size は学習済みモデルと同じ値を設定してください。

追加学習（データセットの選択）

■ 知識ライブラリのデータ管理で次のような操作をしている場合、追加学習ウィザードのデータセットの選択画面で「[ベース入力イメージデータ] が存在しません」とメッセージが表示され、データが見えなくなることがあります。

- 追加学習で使用するベースイメージデータを削除
- 追加学習で使用するベースイメージデータ種別を「教師データ」から変更

データセット管理で次のような操作をしている場合、追加学習ウィザードのデータセットの選択画面で「[ベースデータセット] が存在しません」とメッセージが表示されデータが見えなくなることがあります。

- 追加学習で使用するベースデータセットを削除

【回避方法】

追加学習の実行前に次のことを確認し、前述のような操作を行わないでください。

- ベースデータセットが存在していること
- ベース入力イメージデータが存在していること
- ベース入力イメージデータのデータ種別が「教師データ」であること

NN 最適化

■ NN 最適化では次の機能との連携はできませんが、ウィザードを起動できてしまいます。
- 追加学習
- 認識

【回避方法】

NN 最適化では追加学習、認識と連携できないため、これらの操作を行わないでください。

■ NN 最適化は、ハイパーパラメーターの評価にテストロスを使用しています。
train_val.prototxt では、Loss レイヤーの top キーの値は、次のどちらかの名前にしてください。

- 1) 'loss'
- 2) 'loss1/loss1' と 'loss2/loss1' と 'loss3/loss3'

補足：

- 2) の場合、3 つのレイヤーの平均値をテストロスとして使用します。
- 「種別：スタンダード」の定義の場合は、上記の形式に沿っています。
 - LeNet および AlexNet は、1) の形式です。
 - GoogleNet は、2) の形式です。

対話型学習

- /workspace の残り容量が 50GB 未満の状態に対話型コンテナを停止してもコンテナイメージは保存できません。
対話型コンテナ停止時にコンテナイメージを保存するために、/workspace の残り容量が 50GB 以上必要です。残り容量が不足している場合、警告ダイアログが表示されますのでファイルの削除などを実施し、空き容量を確保してください。
空き容量を確保せずに停止処理を続行した場合、以下のような影響があります。

【お客様影響】

- 保存済みのコンテナを再開して使用中であった場合：
次回再開時、最後の停止時のコンテナイメージは起動されず、1 世代前のコンテナイメージが起動されます。
- 新規作成で起動したコンテナを使用中であった場合：
コンテナイメージは保存されていないため、状態はエラー表示となり再開することはできません。

- 対話型の Docker コンテナ内の大量のファイル操作で、永続的に対話型学習が再開できなくなる場合があります。
現象発生時は、対話型学習の再開で「エラー（再開可）」となり、その後、再開を何度行っても AB32621 のエラーとなります。

再開時のエラー例：

```
INFO: AB32621: jupyter start denied. (in-use_GPUs: xxx) MCODE='01030000-0000-00000000'  
Some reserved GPUs in use.
```

上記の事象となった場合は、当社サポート窓口にご連絡してください。

【回避方法】

対話型は以下のように使用してください。

- 大容量のパッケージをインストールする必要がある場合、ダウンロードファイルはストレージ（NAS）上に配置する
- データセットやアプリケーションのログなどは、ストレージ（NAS）上に配置する

対話型の Docker コンテナ内のファイル操作で、ファイルシステムに対する更新量に応じて対話型学習の再開に時間がかかるようになります。再開に時間のかかる対話型 ID を運用した場合に、本現象が発生することがあります。

なお、すでに Docker コンテナ内で大量のファイル操作を行ったことで、再開に時間がかかっている場合、時間を短縮する方法はありません。

ファイルを削除するなどの対処を行っても、再開時間を短くすることはできません。

すでに該当するファイル操作を行ったことで対話型学習の停止・再開に時間がかかる場合、対話型環境を新規作成し、ストレージ（NAS）に資産を置く運用にて再構築する必要があります。

対話型学習（時系列データ解析）

- 対話型（種別 :Time Series Data Analysis）で入力ファイルのエラーに対し、メッセージなどが出力されないか、Segmentation fault となり、対処方法が特定できない場合があります。

【対処方法】

train.py で原因が特定できないエラーとなった場合は、次の事象に該当しないかを確認してください。

- データセットの系列名（3 列目）に指定した値が、データ間隔チェック用閾値ファイルに存在しない場合エラー終了します。
- データセットの時刻（4 列目）に時刻を指定するとき、時刻間隔が狭すぎる場合はエラー終了します。
- データセットの同一の ID で有効な系列データが 3 未満の場合 Segmentation fault となります。
- データセットの ID の種類（数）がパラメーターファイルの "batchsize_div" より小さい場合、エラー終了します。

学習（Caffe）

- Caffe による学習の中断時に、Caffe プロセスがハングアップすることがあります。このとき中断時のスナップショットは保存されません。

学習種別により次のような影響があります。

- バッチ型学習：中断ボタンによるスナップショットが保存されない場合があります。この場合、再開時には一つ前のスナップショットから再開されます。
- NN 最適化：各タスク終了時にスナップショットを採取しますが、中断ボタン操作時に学習中状態であった場合にスナップショットが保存されない場合があります。
- 対話型学習：Ctrl-C にてプロセスが停止しない場合があり、Ctrl-C 時にスナップショットを採取するオプションを指定していてもスナップショットが保存できない場合があります。

【回避方法】

この事象を回避する方法はありませんが、以下のようにスナップショット採取間隔やタイミングを調整することにより損失を軽減することができます。

- バッチ型学習：スナップショットをエポック毎に採取してください。エポック毎の採取が採取間隔設定 (Snapshot interval) の初期値です。
- NN 最適化：サマリより、誤答率の収束状況をみて中断を行ってください。
- 対話型学習：caffe コマンドの出力より中断タイミングを調整してください。

エッジ画像管理

- 「エッジ画像管理」において、登録される画像が多くなるに伴い、表示にかかる時間が増加します。
1,000 件の画像が表示されている場合、約 10 秒を要します。

【回避方法】

登録済み画像を適宜削除してください。

エッジ推論 SDK (Android, iOS)

- 障害修正を取り込んだ最新の SDK を利用するためには、SDK ダウンロード画面から SDK をダウンロードし、御利用中のライブラリを最新のライブラリに置き換えてください。
障害修正内容については、SDK に同梱されているガイドを御確認ください。

モデル詳細ジョブ情報

- 1 つの学習モデルに対して複数のジョブが実行された場合、最後に実行されたジョブの情報だけ表示されます。

【対処方法】

学習モデルに紐づくジョブを確認する場合、ジョブ状況画面で御確認ください。

モデル詳細画面

- モデル詳細画面でスナップショットを選択し「ダウンロード」ボタンを押下、またはモデル詳細 (NN 最適化) 画面で実行結果のモデルを選択し「ダウンロード」ボタンを押下した直後に、別画面に移動するとダウンロードに失敗します。

【対処方法】

ダウンロードボタン押下後保存用ダイアログが表示されるまでは、以下の操作を行わないでください。

- 他画面遷移
- ダウンロードボタン連続押下

スナップショットダウンロード

- ダウンロード不可の結果をダウンロードしようとする、通常のエラー出力ではなく、エラーメッセージを含むテキストがブラウザに表示されます。

【回避方法】

ストレージ容量不足でスナップショットが保存されない状態で、スナップショットをダウンロードすると、エラーが表示されます。

学習を行う前に、十分な空き容量があることを確認してください。

知識ライブラリ

- 知識ライブラリのデータのマージウィザードでデータのマージを行い、マージ中にストレージの空き容量が不足した場合、データマージ状況画面のマージ状態に「マージ失敗（内部エラー）」が表示されます。

【回避方法】

ストレージの空き容量を増やしたあと、再度データのマージを行ってください。
データのマージと同時に知識ライブラリにデータを登録する場合は、ストレージの空き容量に注意してください。

知識ライブラリ（対話型データ登録）

- 知識ライブラリの対話型データ登録には時間がかかります。登録中に知識ライブラリの機能を実行した場合、登録が完了するまで実行待ち状態になります。
知識ライブラリの機能を実行する際は、対話型データ登録状態画面にて登録中のデータがないか確認の上、実行を行ってください。
なお、登録中のデータが無い場合でも、実行した機能が完了しない場合は、他テナントで登録が行われている可能性がありますので、しばらくお待ちください。

対話型データ登録にかかる時間の目安

- 1GB：約 5 分
- 10GB：約 45 分
- 100GB：約 7.5 時間
- 700GB：約 2 日

FUJITSU AI ソリューション
Zinrai ディープラーニング システム
制限事項・注意事項

発行日 2018 年 10 月
Copyright 2018 FUJITSU LIMITED

- 本書の内容は、改善のため事前連絡なしに変更することがあります。
- 本書の無断転載を禁じます。