

**Dynamic Enterprise AI Solution** 

# <AlRead White Paper> AlRead 手書きmultiモデル用 学習ツール 使用手順書

Confidential

2023年6月 (AWP-002-10)



AIとローコード開発を活用し 世の中をより良い世界に



- 1. はじめに
- 2. 学習ツールのインストール
- 3. 教師データの出力設定
- 4. AIRead FormEditorで手書きモデルを指定
- 5. AIRead Viewerの起動
- 6. AIRead Viewerで修正と保存
- 7. 追加学習の実行
- 8.動作の確認





- ■本書は学習ツール「AlRead training\_tool」の利用手順書を記載した文書です。
- ■本機能はオンプレス版の座標指定OCRでのみ利用が可能です。
- 本ツールで学習可能なモデルは、以下となります。
  - 数値・電話番号 multi\_number
  - 数字 multi\_numeric
  - カタカナ multi\_katakana
     ※2023年6月(AIRead Ver. 4.2.0)時点
- AIReadは、Ver. 3.2.0 以降のバージョンが前提となります。

■ AIReadに手書きモデルがインストールされている前提となります。



# 1. はじめに

■ 「AIRead」と「AIRead training\_tool」の関係、学習の機能と流れは下図の通りです。



/Irise Innovation, Inc.

2. 学習ツールのインストール(1/3)

「AIRead training\_tool」のインストールを行います。

## ■ 手順:

① 「AlRead training\_tool」のモジュールは、以下のサイトよりダウンロードができます。

#### AlRead Downloadサイト (ログイン画面)

URL <u>https://airead.tech/download-list</u>

最新バージョンの「AIRead training\_tool」をダウンロードします。

#### AIRead training\_tool

Version	Release Date	Note
AIRead training tool v1.0.3	2022/05/23	最新バージョン

ファイル名 : training\_tool.zip ファイルサイズ:443 MB バージョン : Ver. 1.0.3

2. 学習ツールのインストール(2/3)

■ 手順:(続き)

③ ダウンロードしたモジュール (zipファイル) を解凍します。





2. 学習ツールのインストール(3/3)

# ■ 手順:(続き)

④ AlReadインストールフォルダ※配下に「training\_tool」フォルダを作成し、解凍後のフォルダ・ ファイルを「training\_tool」フォルダに保存します。

<ul> <li>▶   </li> <li>▶ =   AlRead</li> <li>ファイル ホーム 共有</li> </ul>	表示			100 675	
$\leftarrow \rightarrow \sim \uparrow \blacksquare \rightarrow PC$	> Windows (C:) > AlRead		~ U	,○ AIReadの検索	
クイック アクセス	名前	更新日時	種類	サイズ	^
<ul> <li>OneDrive - Personal</li> </ul>	AIRead_conf AIRead conf kw	2021/12/15 16:17 2021/12/15 17:23	ファイル フォルダー ファイル フォルダー		
PC	aut ClassificDir	2021/12/15 16:13	ファイル フォルダー		
🥩 ネットワーク	ControlPanel	2021/12/15 17:45	ファイル フォルダー		
	debug	2021/12/15 16:47 2021/12/15 16:47	ファイル フォルター ファイル フォルダー		
	Failed FormEditor	2021/12/15 16:03 2021/12/15 18:55	ファイル フォルター ファイル フォルダー		
	IDELibrary input	2021/12/15 17:24 2021/12/15 18:58	ファイル フォルダー ファイル フォルダー		
	📕 java 🤰 logs	2021/12/15 16:03 2021/12/15 16:46	ファイル フォルダー ファイル フォルダー		
	output RuleEditor	2021/12/15 16:19 2021/12/15 17:24	ファイル フォルダー ファイル フォルダー		
	sample SortingEditor	2021/12/15 16:03 2021/12/15 16:03	ファイル フォルダー ファイル フォルダー		
	success	2021/12/15 16:47	ファイルフォルダー		手動で「training_tool」フォル
	training_data	2021/12/15 16:10	ファイル フォルダー		作成し解凍後のフォルダ・ファ
		2021/12/15 16:14	ファイルフォルダー		をほそしよう
	Viewer AlRead_setting.ini	2021/12/15 16:03 2021/12/15 16:46	ファイル フォルダー INI ファイル	3 КВ	
	AlRead_setting_kw.ini	2021/12/15 18:25	INI ファイル	3 KB	~ ·

※AIReadインストールフォルダは、標準は"C:¥AIRead"です

# 3. 教師データの出力設定(1/2)

■ AIRead ControlPanelから教師データの出力設定を行います。

# ■ 手順:

 Windowsの「スタート」から 「AIRead ControlPanel」を起動します。



# ② [学習]タブを選択します。

in control and of the	mises								
設定									Lang : Ji
771N	基本設定	モード別設定	而像処理	信别律处理設定	文字認識調整	ログ出力	ŦR		
☑ 教師デー	タを作成する								
教師データ作	■成フォルダ ①	_/training_da	ta/						
学習するモラ	ルの名称 🛈	multi_numb	er -						
				_	-				
C:¥AlRead¥AlRead	setting_kw.ini				保存				

3. 教師データの出力設定(2/2)

手順: (続き)

③ 「教師データを作成する」にチェックをします。

④「教師データ作成フォルダ」を指定します。※指定が無ければ変更は不要です

☑ 教師データを作成する

教師データ作成フォルダ ①



**Confidential** 

⑤「学習するモデルの名称」に、学習させたい モデル名を入力します。 複数のモデルを学習させる場合、モデル名を (カンマ)で区切ります。

指定できるモデル:

言語	モデル名
数字	multi_numeric
数値・電話番号	multi_number
カタカナ	multi_katakana



 教師データが自動作成されるモデルは「3.教師データの出力設定」で指定したモデルを使用する 項目のみです。

■ 手順:

① AlRead FormEditorで、自動生成をさせたい項目に所定の手書きモデルを指定します。



Arise Innovation, Inc.

Copyright © Arise Innovation, Inc. All Rights Reserved.

Confidential

■ 教師データの自動作成・蓄積はAIRead Viewerで誤読個所を修正する必要があります。

# ■ 手順:

① AlRead Viewerをarex ファイルから実行する場合

- ▶ 「共通設定ファイルパス」に「3. 教師データの出力設定」で設定したファイルを指定します
- ▶ 「AIRead Viewerの起動」に"起動する"を指定します

AIRead ShortcutCreator	on Premises v 4.2.0.0		– 🗆 X
新規作成	起動条件の設定		
開く	読み取り方式の必須	座標指定OCR    ▼	- 3. 教師デーダの出力設定」 で設定した共通設定
保存			ファイルを指定します
名前を付けて保存…	Form ID <mark>①</mark> 必須	order	
	共通設定ファイルパス① <mark>必須</mark>	C:¥AIRead¥AIRead_setting.ini	ファイルの選択
	AIRead Viewerの起動 <mark>①</mark> 必須	◉ 起動する ○ 起動しない	
	読取画像の格納フォルダ <mark>①</mark> 任意	"起動する"にチェック します	フォルダーの選択
			Lang Japanese 💌

- 手順: (続き)
  - ② コマンド (run\_with\_viewer.bat) から実行する場合
    - ▶ "--settingfile"で指定するファイルに「3. 教師データの出力設定」で設定した共通設定ファイルを指定します
    - run\_with\_viewer.bat



- ③ 「AIRead FormEditor」から実行する場合
  - ▶ 左下のメニュー「共通設定ファイル」に「3. 教師データの出力設定」で設定した共通設定ファイルを指定します



## AIRead ViewerでOCRの結果を確認します。

a AlRead Viewer	キャンセル	保存		ファイル名 FormID ステータス	- ロ × :AIRead手書きサンプル帳景 :sample :読取完了
				処理日時	:2023/06/26 15:33:09
			項目名	值	信頼値
			申請日	令和3年2月1日	99
	· 陳式 2		申出人		95
	↑和 3 年 2 月   日 市民税・県民税還付金の振込口座申請書		郵便番号1	104	99
	(現先)○○ 市長		郵便番号2	0054	99
	上記の題付金を下記の口嫌に振り込み願います。		住所	東京都中央区勝どき3-13-1フォアフロン	ントタワー212隆 98
	电出入 岚 行豆 太郎		口座名義力	アライズタロウ	99
	※必ず締税資券者ご未人名浪の口座をご指定ください。		口座名義		98
	ft 所東京都中央区勝じき3-13~1		電話番号	(03)5534-9967	99
	7ヵ17ロンドダイーン「は福 マリガナ アライス・9 ロウ		金融機関J-	0001	99
	四將4歲 崴伊豆太郎		金融機関名	みずほ	99
	□2018年7 (03)5534 - 4467 □-F 金融機関ラーF 0001 友密・芯舗コーF 112		金融機関	銀行	99
			支店コード	112	99
			支店名	大手町	98
	□ 第二十二日 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		本支店	支店	100
	遵付予定金額 ¥ <u>24,800</u> 管理番号 <u>AR 20210201</u> (加速AR 20210201		口座種別	普通	100
	※この書類は AIRead 手書き検証用のサンプル帳景です	•	口座番号	9876543	99
	202101		還付予定金	24,800	99
		<b>—</b>	Andre 1000 107 100		×
					確定
					言語: Japanese ▼

- AIRead Viewerで誤読個所を修正・保存します。
- 手順:
  - 誤読個所を修正します。









\*本書の説明用にOCR結果を加工しています

/Irise Innovation, Inc.

● 手順: (続き)

② AlRead Viewer画面上部の[保存]ボタンを押下します。



③ 所定のフォルダ (C:¥AlRead¥training\_tool) に教師データが自動作成されます。



#### <u>モデルフォルダ</u>

#### <u>教師データと正解情報</u>

PC > OS (C:) > AlRead > training_data > multi_katakana				
名前	日付時刻	種類	サイズ	
<ul> <li>20230626_153309_AlRead手書きサンプル帳票(記入)_0.png_5.jpg</li> <li>20230626_153309_AlRead手書きサンプル帳票(記入)_0.tsv</li> </ul>	2023/06/26 15:34 2023/06/26 15:39	JPG ファイル TSV ファイル	14 KB 1 KB	

PC > OS (C:) > AlRead > training_data > multi_number				
名前	日付時刻	種類	サイズ	
🗋 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_7.jpg	2023/06/26 15:34	JPG ファイル	18 KB	
📄 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_16.jpg	2023/06/26 15:34	JPG ファイル	9 KB	
20230626_153309_AIRead手書きサンプル帳票(記入)_0.tsv	2023/06/26 15:39	TSV ファイル	1 KB	

PC > OS (C:) > AlRead > training_data > multi_numeric			
名前	日付時刻	種類	サイズ
📄 20230626_153309_AlRead手書きサンプル帳票(記入)_0.png_2.jpg	2023/06/26 15:34	JPG ファイル	5 KB
🗋 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_3.jpg	2023/06/26 15:34	JPG ファイル	7 KB
📄 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_8.jpg	2023/06/26 15:34	JPG ファイル	8 KB
📄 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_11.jpg	2023/06/26 15:34	JPG ファイル	5 KB
📄 20230626_153309_AIRead手書きサンプル帳票(記入)_0.png_15.jpg	2023/06/26 15:34	JPG ファイル	14 KB
20230626_153309_AIRead手書きサンプル帳票(記入)_0.tsv	2023/06/26 15:39	TSV ファイル	1 KB

- 手順: (続き)
  - ④ 教師データの内容は下記の通りです。
    - ▶ 教師データ

切り取られた画像ファイルがJPG形式で保存されます。

#### 20230626\_153309\_AIRead手書きサンプル帳票(記入)\_0.png\_5.jpg

アライズタロウ

▶ 正解データ

誤読文字(修正前)、正解文字(修正後)、該当教師データ(画像ファイル保存場所)の情報が、タブ区切りテキスト ファイルで保存されます。

#### 20230626\_153309\_AIRead手書きサンプル帳票(記入)\_0.tsv

🖺 C:¥AlRead¥training_data¥multi_katakana¥20230626_153309_AlRead手書きサンプル帳票(記入)_0.tsv(更新) - sakura 2.4.1.2849	_		×
ファイル( <u>F</u> ) 編集( <u>E</u> ) 変換( <u>C</u> ) 検索( <u>S</u> ) ツール( <u>I</u> ) 設定( <u>O</u> ) ウィンドウ( <u>W</u> ) ヘルプ( <u>H</u> )			
「 ☞ マ   編 冊   局 ♥ 命 端 炎 湯 **! 願 閉 料   画			
	0.png	5. jpg	<del>~</del> ^
[EOF]			
			>
」 1行 1桁 CRLF U+30DE UTF	F-8	REC	挿入





### ■ 専用のコマンドを実行し教師データを用いた追加学習を行います。

#### ■ 手順:

- エクスプローラーから「AlRead training\_tool」を保存したフォルダ(C:¥AlRead¥training\_tool)を開きます。
- ② 追加学習をさせたいモデル名がファイル名に含まれたバッチファイルを実行します。

✓ ↑ ► > PC > OS (C:) > AlRead > training	_tool			学習中の画面です。 実行するPCのスペック、教師
名前	更新日時	種類	サイズ	データの量により完了時間は異 なります(数分〜数時間)。
📙 base-models	2023/06/26 15:29	ファイル フォル	∝ C:¥Windows¥system32¥cmd.exe om version 0.23.1 when using version 0	.22.2.post1. This might lead to breaking code or invalid results. Use at your own
📙 dist	2023/06/26 15:29	ファイル フォル	risk. baded model encoder ['O''I''2''3' ppying C:¥AlRead¥training_data/multi_n puning C:¥AlRead¥training_data(multi_n	'4' '5' '6' '7' '8' '9'] umeric%20230626 [53309_AlRead手書きサンブル帳票(記入) 0.png_2.jpg umeric%20230626 [50309_AlRead手書きサンブル帳票(記入) 0.png_2.jpg
📙 labels	2023/06/26 15:32	ファイル フォル	ppying c.#Alkeau#training_uata/mutri_n bad model ARNING:tensorflow:From tensorflow_core	unier 10#20200020_100009_Airead+音とアフル版系(近人)_0.phib_0.jpp ¥python¥opsYreeSource_variable_ops.py:1630: calling BaseResourceVariableinit riable_ons) with constraint is denrecated and will be removed in a future_version
🔄 AlRead_TrainHW_setting_multi_katakana.ini	2022/03/01 18:33	構成設定	nstructions for updating: f using Keras pass * constraint argume	nta to lavers.
🔄 AlRead_TrainHW_setting_multi_number.ini	2022/03/01 18:33	構成設定	ARNING:tensorflow:From tensorflow core on.ops.init_ops) with dtype is depreca nstructions for updating:	¥python¥ops¥init_ops.py:97: calling VarianceScalinginit (from tensorflow.pyt ted and will be removed in a future version.
🔄 AlRead_TrainHW_setting_multi_numeric.ini	2022/03/01 17:47	構成設定	all initializer instance with the dtyp ARNING:tensorflow:From tensorflow_core s.init_ops) with dtype is deprecated a	e argument instead of passing it to the constructor ¥python¥ops¥init_ops.py:97: calling Orthogonalinit (from tensorflow.python.o nd will be removed in a future version.
💿 train_multi_katakana.bat	2022/03/01 17:47	Windows パ	nstructions for updating: all initializer instance with the dtyp ARNING:tensorflow:From tensorflow_core	e argument instead of passing it to the constructor ¥python¥ops¥init_ops.py:97: calling Zerosinit(from tensorflow.python.ops.in
💿 train_multi_number.bat	2022/03/01 17:56	Windows パッ	t <u>ops) with dtype is deprecated</u> and wi istructions for updating: all initializer instance with the dtyp MDNING tensorflow:From tensorflow or o	n be removed in a ruture version. e argument instead of passing it to the constructor #vubnompowerse.com pv:1475: where (from tensorflow python one array one) is da
Train_multi_numeric.bat	2022/03/01 17:48	Windows /(	recated and will be removed in a futur nstructions for updating: se tf.where in 2.0, which has the same 023-06-27 09:21:16.810700: I tensorflo TensorFlow binary was not compiled to	e version. broadcast rule as np.where w/core/platform/cpu_feature_guard.cc:142] Your CPU supports instructions that thi use: AVX2

# 7. 追加学習の実行

- 手順: (続き)
  - ③ バッチの実行が完了すると、「AlRead training\_tool」を保存したフォルダ (C:¥AlRead¥training\_tool) 配下に下記のフォルダが作成されます。

≻ data	• • •	学習に使用した画像ファイルを保存するフォルダ
≻ logs	• • •	学習結果のログを保存するフォルダ
≻ outputs	• • •	追加学習後のモデルが出力されるフォルダ
➤ test_result	• • •	学習条件の情報が保存されたフォルダ

<b></b> 前	更新日時	種類	サイズ	【注意】  「autputa」「toot rogult」フェルグは営羽
base-models	2023/06/26 15:29	ファイル フォルダー		「「outputs」「test_result」ノオルメは子首 が成功した場合にのみ作成されます。
data	2023/06/27 9:10	ファイル フォルダー		学習が失敗する場合、以下を確認ください。
dist	2023/06/26 15:29	ファイル フォルダー		
labels	2023/06/26 15:32	ファイル フォルダー		 - 教師データが 2件未満
logs	2023/06/27 9:10	ファイル フォルダー		学習には 1つのモデルに対して最低 2件の教師
outputs	2023/06/27 9:10	ファイルフォルダー		データ(調読を修正したデータ)が必要です
test_result	2023/06/27 9:10	ファイル フォルダー		
${\sf AIRead\_TrainHW\_setting\_multi\_katakana.ini}$	2022/03/01 18:33	構成設定	1 KB	
AIRead_TrainHW_setting_multi_number.ini	2022/03/01 18:33	構成設定	1 KB	bh MM an AlPoodTraining log / / 1/50
${\sf AIRead\_TrainHW\_setting\_multi\_numeric.ini}$	2022/03/01 17:47	構成設定	1 KB	IIII_WW_SS-AIReau Training.log」に以下の
test_result.txt	2023/06/27 9:10	TXT ファイル	1 KB	ノダウセーシが出力されより。
train_multi_katakana.bat	2022/03/01 17:47	Windows バッチ ファ	1 KB	
train_multi_number.bat	2022/03/01 17:56	Windows バッチ ファ	1 KB	Exception: At lease 2 image are required
train multi numoric hat	2022/03/01 17:48	Windows バッチ ファ	1 KB	

//rise Innovation, Inc.

# 7. 追加学習の実行



# ■ 手順:(続き)

 ④「outputs」フォルダ(C:¥AlRead¥training\_tool¥outputs)配下に作成された日付時刻のフォルダが 作成され、その配下にモデル名のフォルダが作成されます。

✓ ↑	ad > training_tool > outputs >			> training_tool > outputs > 2	02306270921 >	
名前	更新日時	種類	名前	更新日時	種類	サイズ
202306270921	2023/06/27 9:21	ファイル フォルダー	Multi_numeric	2023/06/27 9:21	ファイル フォルダー	
			202306270921_multi_numeric_mod	del.json 2023/06/27 9:21	JSON Document	48 KB
			202306270921_multi_numeric_wei	ight.h5 2023/06/27 9:21	H5 ファイル	80,661 KB

 5 モデル名のフォルダごと、AlReadのモデル格納フォルダ(C:¥AlRead¥tessdata)にコピーします。 この際、既に存在するモデルフォルダはリネーム等で保存してください。

↑ I > PC > OS (C:) > AlRead > training_tool > outputs > 202306270921 >							
名前	更新日時	種類	サイズ		更新日時	種類	サイズ
	2022/06/27 0.21		י <sup>י</sup> ו	barcode_detector	2023/06/16 14:49	ファイル フォルダー	
Multi_numeric	2023/06/27 9:21	ファイル フィルター	$\mathbf{V}$	char_segmentation	2023/06/16 14:49	ファイル フォルダー	
202306270921_multi_numeric_model.json	2023/06/27 9:21	JSON Document	48 KB	📕 checkmark	2023/06/16 14:49	ファイル フォルダー	
202306270921_multi_numeric_weight.h5	2023/06/27 9:21	H5 ファイル	80,661 KB	📙 eng	2023/06/16 14:55	ファイル フォルダー	
				📙 jpn	2023/06/16 14:54	ファイル フォルダー	
				📜 jpn_address	2023/06/16 14:54	ファイル フォルダー	
				📜 money	2023/06/16 14:54	ファイル フォルダー	
				📜 multi_katakana	2023/06/16 14:54	ファイル フォルダー	
	既存のモデル	レフォルダをリ	ネーム	📕 multi number	2023/06/16 14:54	ファイル フォルダー	
	保存し作成さ	<b>x</b> れたモデルフ	オルダ	Multi_numeric	2023/06/27 9:44	ファイル フォルダー	
	レーズ なー	1ピーします		multi_numeric_old	2023/06/16 14:54	ファイル フォルダー	
	<u>د</u> –			multi2_jpn	2023/06/16 14:54	ファイル フォルダー	
				multi2_jpn_name	2023/06/16 14:54	ファイル フォルダー	

#### //rise Innovation, Inc.



■ 以上で追加学習とモデルの更新は完了です。実際の帳票を読み取り、結果を確認してください。

## 【注意】追加学習を行ったとしても、必ずしも誤読が無くなるわけではありません。





# //rise Innovation, Inc.

アライズイノベーション株式会社

東京都中央区勝どき3-13-1

Mail : <u>airead\_support@ariseinnovation.co.jp</u>

AIRead Official : <u>https://airead.ai/</u>

AIRead Support : <u>https://airead.tech/</u>

AIRead Manual : <u>https://manual.airead.ai/</u>

/Irise Innovation, Inc.