pyImageUtils for Blue Prism
(Deskew image)

USER GUIDE

Contents

1.   Introduction

This Blue Prism Skill deskews images by using Python (embedded) Web Service. The purpose in fact is simple. If you have any deskewed images like scanned documents this Blue Prism Skill will rotate them automatically to make easier for example the OCR work.

This skill does not work alone. It comes with a separate Web Service developed in Python and uses the deskew library described here https://github.com/sbrunner/deskew. By default, the Blue Prism VBO launch (if needed) this web service on the server side and manages the Web services calls to deskew the requested images. You can naturally change this behavior by just adjusting the run service Action in the vbo.

2.   Prerequisites

To make this skill work you need:

1. Blue Prism 6.9 minimum

2. A working Python 3.6 min environment.

For the Python environment you will have to import/install several libraries:

- numpy
- skimage
- deskew
- cv2
- flask
- jsonpickle

For the Blue Prism environment, you will need these vbo:

- Utility - Environment
- Utility – JSON

3.   Configuration

The sections below describe how to use the Blue Prism skill.

3.1.   Python

First you must install a Python environment. I recommend to install Anaconda (https://www.anaconda.com/) or you can just install Python here : https://www.python.org/ Once you've installed Python you'll need to install additional libraries to make the Web Service work. To do that you can use the pip utility or conda if you're working with anaconda.

These are the required libraries:

- numpy (just install by typing in the command line *$ pip install numpy*)
- skimage
- deskew
- cv2
- flask
- jsonpickle
- pdf2image

Note: a windows command file (**install_python_packages.cmd**) regroups all these libraries installation directives.

One all these packages have been successfully installed you can start by copying the files (into the Github directory) locally: ie. into a Blue Prism Windows server folder.

Now, open the **runWsImageUtils.bat** file

@CALL C:\Users\admin\Anaconda3\Scripts\activate.bat C:\Users\admin\Anaconda3
python "C:\BP Assets\services\imageutils\imageutils.py

You need to change 2 things:

1. The anaconda directory to reflect your Python environment (if you use Anaconda distribution).
2. The directory to reflect where you had copied the files previously.

Normally you don't have to change anything in the Python code. I know the exception management and other good developer stuff is not yet done but i would want it simple and easy to adapt and change. So, do not hesitate to make your modifications in there (and share it through GitHub).

3.2.   Blue Prism

Firstly, you have to ensure that libraries are already imported into Blue Prism before importing this skill:
- Utility - Environment
- Utility – JSON

Note: these vbo are by default provided by Blue Prism (by default in the C:\Program Files\Blue Prism Limited\Blue Prism Automate\VBO) and are copied during the installation of the product. However, they are not imported, you must do the import (or checking if they are already installed) manually by yourself before going further.

Once these objects/vbo are imported successfully, you can import the deskew skill:

1. Open the Blue Prism Studio and import the bprelease file. Once you have done that you should have two new assets: a web service and a vbo (object).
2. If you have made some change in the Python code like for example changing the URL (or the port) you may need now to do some change in the Blue Prism Web Service definition. Otherwise do not change anything at this stage.
3. Change the Blue Prism environment variable value (for *pyRunImageUtilsCommandLine*) to refer to the command line path you have copied (in the GitHub bundle this is the *runWsImageUtils.bat* file) in the Blue Prism Environment variable settings.
4. Open the vbo named pyImageUtils
5. save it

Note: The Blue Prism export is compatible for Blue Prism 6.9 and later.

3.3. Last check

The new skill should work now. Just run the deskew action (in the pyImageUtils object) and see the generated image (Cf. https://www.datacorner.fr/bp-deskew/)

4.   Using the Asset

The Visual Business Object **pyImageUtils** contains the following actions:

4.1.   deskew
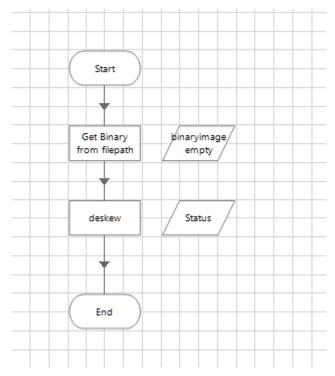
Inputs:

Name	Description	Data Type
image	The input file to deskew in binary format	binary
targetfile	The filename of the target and deskewed file	text

Outputs:

Name	Description	Data Type
Status	Returns Saved if the work has been done successfully	text

Try the provided example (Blue Prism Process : deskew image example) by giving an image as input :

The first action (Get Binary from filepath) is used to get the image from a filepath and converting it as a binary format.

The image below is given as example in the data folder, try the process with :

The second action send the image (binary format) to the deskew service (if this one does not run, it's launched automatically). You should see a command line window opening and also see some trace in there proving the service is working. The deskewed image is then written on the file path given as parameter.

You should see a new file appearing in the destination folder :

4.2. deskewpdf

Inputs:

Name	Description	Data Type
Pdf file	PDF file to deskew	binary
target directory	Filepath to store the deskewed image	text
pages to extract	Specify here the page to extract into the pdf (separated by a comma)	text
prefix	File frefix used to store the deskewed files	text

Outputs:

Name	Description	Data Type
Status	Returns Saved if the work has been done successfully	text

About the pdf management the purpose is the same as the one described for Action deskew. However in this case the source file is not an image but a pdf file. In this case there's a different preparation phase as the service will breakdown the pdf per pages and convert them as image. You have the possibility to specify which page you want to extract.

Try the provided example (Blue Prism Process : deskew pdf example).

4.3.   Run Service

This action is normally called automatically by the deskew action as it launches the web services if this one is not still running.

Inputs:

Name	Description	Data Type
No input		

Outputs:

Name	Description	Data Type
Running	Returns "Service running" if the service if up and ready	text

5.   Support

Please add an issue into Github here : https://github.com/datacorner/blueprism-deskew-skill/issues

You can also add a comment in my website here : https://www.datacorner.fr/bp-deskew/

6. Frequently Asked Questions

Cf. GitHub : https://github.com/datacorner/blueprism-deskew-skill

7. Changelog