



umango

処理の分離手順書

処理の分離

Umango を使用する利点の 1 つは、大量の文書を一度に処理またはスキャンできることです。また、Umango は大量の文書の区切りや各文書の種類を自動的に検出できます。Umango で実行できる分離には 2 つの種類があります。

The screenshot shows the '分離' (Separation) tab in the Umango interface. It is divided into two main sections: 'ドキュメントの分離' (Document Separation) and 'バッチ分離' (Batch Separation).

ドキュメントの分離
大量の原稿を一度にスキャンし、原稿の切れ目を自動的に検出

- ☐ 分離しないで
- ☐ 人工知能
- ☐ 分離シート
- ☐ 空白ページ
- ☐ 最初のページのバーコード
- ☐ 最初のページの OCR テキスト

☐ 毎回ページ数で区切ります。 ページ

☒ 文書の自動分類

☐ 最初のページにアンカー

オプション
精度の信頼性

低い 高い

バッチ分離
一度に複数のバッチをスキャンします。カバーシート「ゾーン」が使用されている場合に便利です。

注: デバイスからスキャンする場合はサポートされていません。

- ☒ 分離しないで
- ☐ 分離シート
- ☐ 自動カバーシート認識

処理の分離

- 1. ドキュメントの分離:** ドキュメント分離は、Umango が大量のドキュメントを一度にスキャンする際にドキュメントの破損を検出するために使用する方法です。この機能により、Umango はドキュメントの種類を検出することもできます。
- 2. バッチの分離:** バッチ分離は、Umango が複数のスキャンを同時に行う際に新しいバッチを検出するために使用する方法です。カバーシートゾーンが使用されている場合に役立ちます。ただし、デバイスからのスキャンではサポートされないことに注意してください。

この機能を使用すると、一度に多くの文書をスキャン/アップロード/Umango に送信するだけで、残りの作業は Umango に任せることができます。

I. ドキュメントの分離

ドキュメント分離とは、大量の書類を一度にスキャンする際に、Umango がドキュメントの区切りを検出するために使用する方法です。

ドキュメントの分離

大量の原稿を一度にスキャンし、原稿の切れ目を自動的に検出

<input type="radio"/> 分離しないで	<input type="radio"/> 毎回ページ数で区切ります。
<input type="radio"/> 人工知能	<input type="text" value="1"/> ページ
<input type="radio"/> 分離シート	<input checked="" type="radio"/> 文書の自動分類
<input type="radio"/> 空白ページ	<input type="radio"/> 最初のページにアンカー
<input type="radio"/> 最初のページのバーコード	
<input type="radio"/> 最初のページの OCR テキスト	

ドキュメントの分離設定

1. 分離しないで

1 つのスキャン／ファイルを 1 つのドキュメントとして扱います。

2. 人工知能

AI を使用してページを解析し、それらが同じドキュメントに属するか、別々のドキュメントとすべきかを判断します。

この方法は、文書の種類によって効果に差が出る場合があります。

ページ番号の記載や、ドキュメント内でページ間に論理的なつながりがある場合、認識精度が向上します。

※テキスト内容が含まれない書類にはこの方法は使用できません。

※AI はクラウド機能として実行されるため、インターネット接続が必要です。

3. 分離シート

Umango のドキュメント分離シートを印刷し、各ドキュメントの間に挟んでスキャンします（最初のドキュメントには不要）。

このシートが新しいドキュメントの検出をトリガーし、スキャンまたはインポート時に自動的に除外され、最終ドキュメントには含まれません。

4. 最初のページのバーコード

スキャンされるドキュメントの 1 ページ目に毎回同じバーコードが含まれている場合、それを検出して分割をトリガーできます。

ゾーンタブで作成したバーコードゾーンを選択し、識別するためのテキストまたは正規表現を入力します。

※分離用にゾーンを使用する場合、ゾーン内の Smart Seek ルールは適用されません。

5. 最初のページの OCR テキスト

スキャンされるドキュメントの 1 ページ目に、毎回同じか類似したテキストが含まれている場合、そのテキストを使用してドキュメントを分割できます。

識別用ゾーンを選択し、識別テキストまたは正規表現を入力します。

※分離用にゾーンを使用する場合、ゾーン内の Smart Seek ルールは適用されません。

6. 毎回ページ数で区切ります。

スキャンされたページ数をカウントし、指定されたページ数に達したら新しいドキュメントを作成します。

(両面スキャンを使用する場合は、裏面も考慮してページ数を設定してください。)

7. 文書の自動分類

各ドキュメントの最初のページが一貫した見た目・レイアウトである場合、自動ドキュメント認識を使用して新しいドキュメントの開始を特定できます。

ジョブで使用するサンプルドキュメントをテンプレートとして登録します。

分割をトリガーするために、認識の信頼度 (Accuracy Confidence) のスライダーでしきい値を調整できます。

8. 最初のページにアンカー

ドキュメントの 1 ページ目に「アンカーポイント」が存在し、かつ信頼度基準を満たした場合に、分割をトリガーできます。

この機能を使うには、ドキュメントの 1 ページ目にアンカーゾーンを設置しておく必要があります。

II. バッチの分離

バッチ分割とは、複数のバッチ（まとめられた文書群）を一度にスキャンする際に、Umango が新しいバッチの開始を検出する方法です。

カバーシートゾーン（バーコードや識別用情報を含むシート）を使用している場合に特に役立ちます。

※注意：デバイス（複合機など）から直接スキャンする場合、この機能はサポートされていません。

バッチ分離

一度に複数のバッチをスキャンします。カバーシート「ゾーン」が使用されている場合に便利です。

注: デバイスからスキャンする場合はサポートされていません。

☒ 分離しないで

☐ 分離シート

☐ 自動カバーシート認識

バッチの分離設定

バッチ分割には、次の3つのオプションがあります。

1. 分離しないで

1回のスキャン／ファイルを1つのバッチとして扱います。

2. 分離シート

Umango のバッチ分離シートを印刷し、各新しいバッチの最初に挟みます（最初のバッチには不要）。

このシートが新しいバッチの開始を示し、スキャンまたはインポート時に自動的に除去され、最終的なドキュメントには含まれません。

3. 自動カバーシート認識

ゾーンタブにアップロードしたサンプルカバーシートを使い、Umango がカバーシートを自動的に識別してバッチ分割をトリガーします。

分割を実行する前に、認識の信頼度（Accuracy Confidence）のスライダーを調整して、必要な認識レベルを設定できます。