

ドローン低層リモートセンシングと深層学習に基づく水田における  
イヌビエの自動認識に関する研究  
Research on automatic recognition of barnyard grass in paddy fields based on  
drone low-rise remote sensing and deep learning

唐 艶穎  
TANG, Yanying

**概要:** 日本の農業は少子高齢化の影響で、農業従事者の高齢化、後継者不足が問題となっている。持続可能な農業を実現するために、農業の省力化・軽労化を目指すスマート農業が推進されている。農業分野では、米を生産するためのイネの育成において、雑草の除去が非常に手間とコストがかかる。特に、イネによく似ているイヌビエの除去が難しい。除草剤の過剰使用によるイネの生育の妨げ、収穫量の低下など、多くの問題がある。本研究では、イネに擬態し、稲作のサイクルに適応した水田雑草であるイヌビエを研究対象とし、ドローン低層リモートセンシングと深層学習技術 (U-Net セマンティックセグメンテーションモデル) を用いたイヌビエ自動認識プラットフォームを構築し、検証を行った。本プラットフォームはビデオや遠隔監視カメラに直接応用できるという汎用性も兼ね備えている。

**Summary:** Due to the declining birthrate and aging population, Japanese agriculture is facing problems such as the aging of agricultural workers and the shortage of successors. In order to realize sustainable agriculture, smart agriculture is being promoted with the aim of saving labor and reducing labor. In the agricultural field, weed removal is very laborious and costly in growing rice for rice production. In particular, it is difficult to remove barnyard grass, which is very similar to rice. There are many problems such as hindering the growth of rice and reducing the yield due to overuse of herbicides. In this research, barnyard grass, which is a paddy weed that mimics rice and adapts to the rice cultivation cycle, is selected as our research target and an automatic barnyard grass recognition platform using drone low-rise remote sensing and deep learning technology (U-Net semantic segmentation model) is developed and Verified. The platform also has the versatility of being directly applicable to video and remote surveillance cameras.

**キーワード:** イヌビエ、ドローン、深層学習、U-Net セマンティックセグメンテーションモデル、自動認識プラットフォーム

**Keywords:** Barnyard grass, Drone, Deep learning, U-Net Semantic Segmentation Model, Auto-recognition platform

## 1. はじめに

### 1.1 研究背景

日本の農業は少子高齢化の影響で、農業従事者の高齢化、後継者不足が問題となっている[1]。農業分野では、米を生産するためのイネの育成において、雑草の除去が非常に手間とコストがかかる。特に、イネによく似ているイヌビエの除去が難しい。除草剤の過剰使用によるイネの生育の妨げ、収穫量の低下など、多くの問題がある[2]。持続可能な農業を実現するために、農業の省力化・軽労化を目指すスマート農業が推進されている[3]。

イヌビエはイネと形が良く似ており、イネとの競合力が非常に強い雑草である。イヌビエは水田雑草の中でも生育量が大いため、養分や光をよこどりして、イネの生育を妨げ、収穫量を低下させている[4]。穂が出る前のイヌビエは、イネと見分けるのが難しい。三葉期を過ぎると除草剤による防除が難しくなり、イヌビエは成熟すると種が土におちて数年間生きることができる。したがって、翌年以降のイヌビエの数が増えるのを防ぐために

は、完全に成熟する前のイヌビエの除去が必要である。

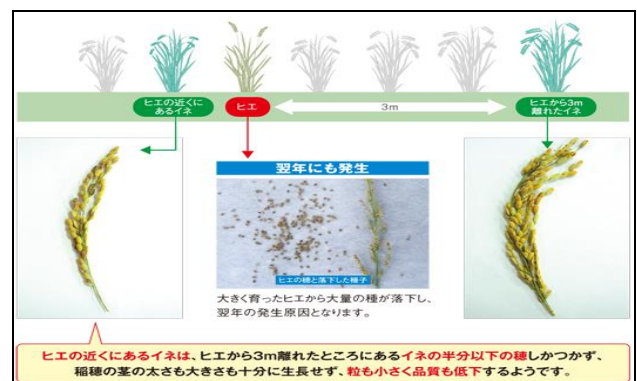


図1 イヌビエがイネの生育に与える影響 出典[4]

現在、日本国内外でAIの画像分析技術を利用して農作物の病虫情報を早期に発見し、対応方法を提示する検証や実際の応用事例が多い。ドローン技術の進歩に伴い、ドローン低層リモートセンシング技術もスマート農業に広く応用されている[3]。しかし、AIとドローン低層リ

モートセンシングを組み合わせる雑草を識別する研究や応用は少ない。

## 1.2 研究目的

本研究では、イネに擬態し、稲作のサイクルに適応した水田雑草であるイヌビエを研究対象とし、ドローン低層リモートセンシングと深層学習技術（U-Net セマンティックセグメンテーションモデル）を用いたイヌビエ自動認識プラットフォームを構築し、実証実験による検証を経て、実用化を目指す。

## 2. 研究用主要技術

機械学習は、成長を続けるデータサイエンス分野の重要な一部である。主に統計的手法を用いてアルゴリズムを訓練し、大量のデータから数値を予測する回帰問題と、データからクラスを分ける分類問題に大別することができる。本研究では大量の画像からイヌビエを判別するため分類問題を題材として機械学習を行う。

セマンティックセグメンテーションは、画像処理やマシンビジョン技術における画像理解の重要な一環であり、AI 分野の重要な分野である。深層学習の発展に伴い、畳み込みニューラルネットワークによるセマンティックセグメンテーション方法は、ネットワークを利用して画像の特徴を自動的に学習し、エンドツーエンドの分類学習を行うことである。セマンティックセグレーションとは、画像内の各画素点を分類し、各点のカテゴリを決定し、領域を分割することである。画像分類とは異なり、セマンティックセグメンテーションはピクセルレベルであり、画像分類は画像レベルである[5]。CNN は画像全体のカテゴリしか識別できず、各画素点のカテゴリは識別できないため、この完全接続方法は画像内の各画素点の分類には使用できない。この問題に対して、一般的な物体認識用の CNN をセマンティックセグメンテーション用の FCN に改良することが提案された[6]。

本研究で使用した U-Net モデルは FCN の一種である。

## 3. イヌビエ自動認識プラットフォームの開発

### 3.1 開発環境の構築

本研究では、まず、イヌビエ自動認識プラットフォームの開発に必要な開発環境を構築する必要がある。本研究において Python を基軸としたプログラミング開発を行う。使用する機械学習フレームワークの多くがオープンソースライブラリとなっており、誰でも無料で利用することができる。その中素早く習得しやすくコードがシンプルで柔軟で、Python で利用できる PyTorch と呼ばれるフレームワークを採用する。本研究では、マイクロソフト社が公開している Visual Studio Code (VS Code) エディタを使用した。Python 言語の開発環境を構築できる。機械学習を行う際にかなり強力なマシンパワーが必要となる。筆者のマシンはプロセッサが I9・11900K、メモリが 32GB となっている。ライブラリは OpenCV、PIL、Numpy である。ローカルで PyTorch をインストールし、GPU は

NVIDIA RTX 3090 であり、高負荷な計算を外部に依存しながら行うことができるようになっており、より大量データを扱う環境が揃っていると言える。

### 3.2 U-Net セマンティックセグメンテーションネットワークの構築

本研究では U-net を用いてドローン画像のイヌビエ自動認識にセマンティックセグメンテーションを行う。

U-Net は Olaf からによって生物医学のために開発されたセマンティックセグメンテーション用のモデルである[7]。U-Net ネットワークは FCN ネットワークを基盤としており、全体的に FCN と考え方が非常に似ている。U-Net は FCN や SegNet をベースにした改良法と見ることができ、FCN の全畳み込み、逆畳み込みアップサンプリング、越段接続の方法を採用し、SegNet の encoder-decoder 構造を採用している。

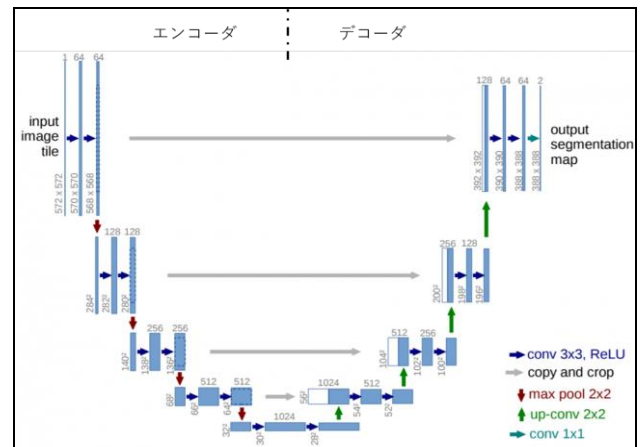


図2 U-Net の構造 出典[7]

図2に U-Net ネットワークの構造を示す。エンコーダやデコーダの2つに分け、大きなU文字に似ている。まず、畳み込みやプールの化を行う。そして逆畳み込みでアップサンプリングを行い、crop の前の下位 feature map で融合を行う。再びサンプリングする出力の feature map が得られるまでこれを繰り返す、softmax を経て出力分割結果が得られる。

本研究では、基幹特徴抽出ネットワークの構築→強化特徴抽出ネットワークの構築→予測ネットワークの構築の順でセマンティックセグメンテーションネットワークを構築した。

## 4 画像データの収集およびデータセットの作成

### 4.1 画像データの収集

イヌビエを自動認識する機械学習モデルを作成する際に大量の画像データが必要になる。本研究ではイヌビエ認識を確実にするため、ドローンで実際に画像を収集する方法を用いた。座間市役所農政課の協力のもとで水田で認可をうけたうえで調査を行った。

本研究は前提としてイヌビエが完全に熟す前の時期にかけて行う調査を対象としている。9月に3回を分けて

同じ地点で撮影を行った。データの有効性を保証するため、飛行高度は2メートルに統一されている。撮影した画像は300枚程度で、中でもイヌビエの特徴が強く出ている140枚を選出し、処理を行う。

#### 4.2 データセットの作成

ドローンで撮影した画像を収集した後、学習セット、テストセット、検証セットの3種類のデータセットを作成することが必要である。

##### 4.2.1 画像の分割

コンピュータによってはメモリが小さいため、一枚の画像をそのまま処理することができないため、大きな画像をブロック化して入力し、結果を一枚ずつつなげて学習をすることが必要である。本研究では、スライドウィンドウカットを利用して4000×2250解像度の原画像を400×450解像度の画像50枚に分割した(図3を参照)。

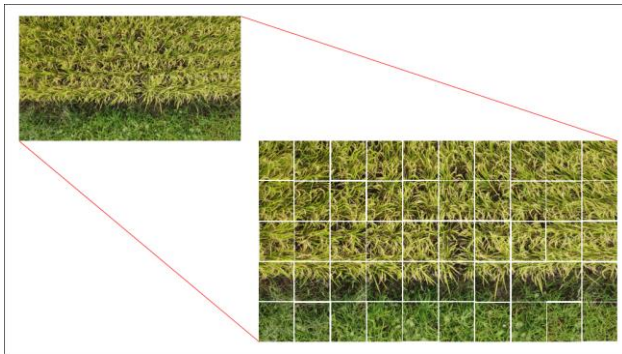


図3 画像の分割結果例

##### 4.2.2 データのアノテーション

アノテーションはAIに教え込ませるためのデータを作る作業のことである。今回は、ピクセルごとにイヌビエと背景に分かれたデータを作る作業である。本研究ではアノテーションツールとしてlabelmeを用い、2000枚アノテーションされた画像を作成し、そのデータを使った(図4)。



図4 アノテーション例

##### 4.2.3 画像データの前処理

より多くのデータセットを獲得し、特徴を充実させ、モデルの過剰適合をふせぐためには、十分なデータセットを用意する必要がある。しかし、実際のタスクでは、十分なデータセットを入手することは難しく、十分なデ

ータを得るために前処理を行う必要がある。本研究のデータの前処理については、回転、左右交換、斜め四角形変化、垂直ひずみ、拡大または縮小、弾性ねじれ、輝度変換の方法を採用した。

図5に示すように、元の画像aを拡張する際に、上述した異なる拡張方法をランダムに組み合わせ、拡張結果(b)を40枚得た。データを前処理した結果、8000枚のデータが得られた。

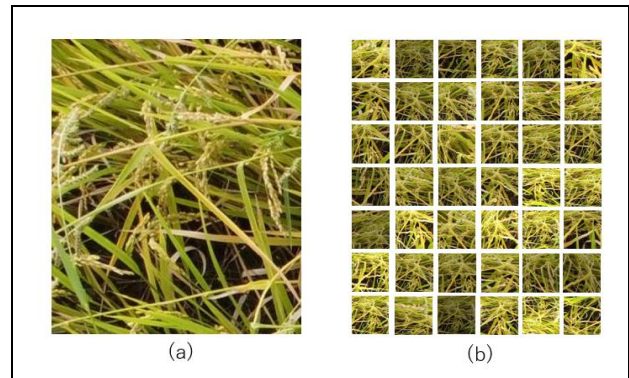


図5 画像の拡張例

##### 4.2.4 データの準備

準備するデータは、集めた画像と分類後の画像の二種類がある。JSONファイルのポリゴンデータを使って、最終的にAIに出力させたい画像を作っている。この分類後の画像はピクセル単位で分類されているものになる。例えば、図6に示すように0が背景、1がイヌビエというように分類ごとに数値で分ける。



図6 準備したデータ例

##### 4.2.5 データセットの区分

入力データや教師データを準備したら、理論的には、データを8:1:1で学習セット、テストセット、検証セットに分ける必要がある。しかし、実際のトレーニングでは、学習セットやテストセットを組み合わせ、コードではデータを9:1で学習セットと検証セットに分けている。学習セットを利用してトレーニングする。

#### 5 トレーニングと検証

##### 5.1 トレーニングと結果分析

セマンティックセグメンテーションに必要なデータセットの作成が完了後、トレーニングを行う。本研究では

最適化手法 Adam を用いる。パラメータ設定：学習率は 0.0001、Batchsize は 128、エポックは 300 となる。

図 7 はトレーニング過程のロスの変移を示したもので、ラウンドごとのトレーニングの loss を記録したものである。

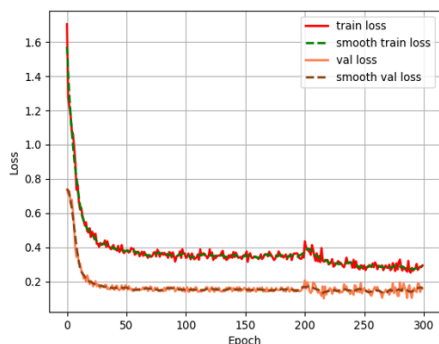


図 7 loss の変化の傾向

このように、トレーニングの初期段階では、loss の低下が早く、トレーニングラウンドが 100 回になると、ほぼ安定したペースで推移していることがわかる。200 回で凍結世代が導入され、一時的に loss が上昇しているのがわかるが、これはバッチサイズが 128 から 64 に変更されたため、250 回ではほぼ安定したペースで推移している。

## 5.2 検証

検証の結果では、クラス平均画素精度 MPA は 88.4 であったため、イヌビエの認識率は 88.4 であったことがわかった。

本モデルの実行後に画像名を入力すれば図 8 のような可視化予測結果が得られる。

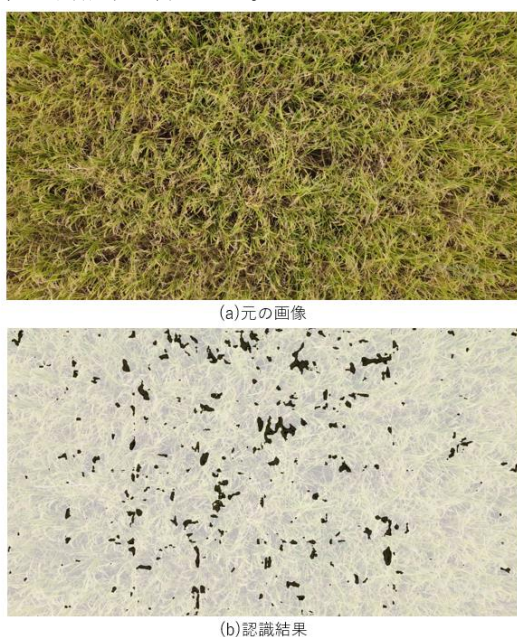


図 8 画像全体イヌビエ認識結果の例

## 5.3 応用例

実用性を高めるために、本研究で開発した予測モデル

は、ビデオや遠隔の監視カメラのイヌビエ自動認識に適用し、予測の過程でイヌビエの面積を自動的に統計し、可視化できる。イヌビエの面積は画素を表している。画素全体に対応する実面積がいくらかがわかると、イヌビエの画素が占める実面積が得られる。



図 9 ビデオに対する自動認識結果例

図 9 に示すのは、本モデルをドローンビデオに実際に適用した自動認識結果の例である。識別画面の左上には、イヌビエがフレームごとに占める画素面積が表示される。

## 6 まとめ

本研究では、イヌビエを自動認識するために、機械学習を用いたセマンティックセグメンテーションモデルを開発し、実証実験による検証を行った。本プラットフォームはビデオや遠隔監視カメラに直接応用できるという汎用性も兼ね備えている。また、イヌビエの画像を収集するには、大量の高品質データが必要である。本研究では、ドローンを利用して収集した画像がセマンティックセグメンテーション自動認識に適した高品質の画像であることが実証され、認識精度は 88.44 と高い。今後、イヌビエの他の成長期および他の種類の雑草におけるセマンティックセグメンテーションモデルを開発すれば、より実用性の高いプラットフォームの構築が可能である。

## 参考文献

- [1] 少子高齢化時代の農業プロセス
- [2] 水田への除草剤散布の注意点とは？使用が適した除草剤は？
- [3] ドローンで農水省と実現する「スマート農業」
- [4] シンジェンタ 稲の減収を招くヒエー稲の生育に与える悪影響
- [5] セマンティックセグメンテーション-画像処理およびコンピュータビジョン
- [6] Evan Shelhamer, Jonathan Long, Trevor Darrell, 「Fully convolutional networks for semantic segmentation」 2015
- [7] Olaf Ronneberger, Philipp Fischer, Thomas Brox, 「U-Net: Convolutional Networks for Biomedical Image Segmentation」 2015