遺伝要因解明のための楕円フーリエ解析を 用いた圃場のダイズ根の形状解析

須藤 友貴^{1,a)} 森山 一斗² 内海 ゆづ子^{1,b)} 松村 篤^{3,c)} 岩村 雅一^{1,d)}

概要:植物の根の形状は植物の栄養獲得戦略に大きく影響を与えるため,生育環境に合う形状の根をもつ植物の育種を実現することが望まれている.効率的な育種は遺伝子操作であるが,根形状の関連遺伝子の特定が必要である.近年,ゲノムワイド関連解析 (Genome-Wide Association Studies : GWAS) で遺伝子の特定が容易になったが,GWASの適用には根の形状を定量化する必要がある.根形状を定量化した既存の研究は屋内栽培のダイズを用いており,通常,ダイズが生育される圃場との環境差により根形状が異なるため,解析手法や結果は実用的でない.そこで,本研究では圃場で栽培されたダイズ根画像から根の概形を抽出し,楕円フーリエ解析を用いて形状の定量化を行う.本研究で得たダイズ根の形状の特徴量をゲノムワイド関連解析で遺伝子との関連を調べた結果,形状の特徴量と関連する遺伝子の変異が一つ検出された.

1. はじめに

植物の根は,養分の吸収や植物と微生物の相互作用の促進など,植物にとって重要な役割を果たす器官である.根の形状に着目すると,地中深くに伸びた根は窒素を吸収しやすく,地中の浅い部分に広がった根はリンを吸収しやすいことが知られている[1].このように,根の形状の違いは植物の栄養獲得戦略に大きく影響を与える.そのため,収量を効率的に増やすためには,生育環境に合う形状の根をもつ植物の育種が重要となる.

本研究で扱う根の形状のような、生物の表現型を変化さ せるには、遺伝子操作が最も効率的である.そして、遺伝 子操作によって表現型を変化させるには、目的の表現型に 関連する遺伝子型を特定する必要がある.表現型と遺伝子 型の関連を解析する方法の一つとして、ゲノムワイド関連 解析 (Genome-Wide Association Studies:GWAS)[2] があ る.この手法は、表現型を数値で表現したものと、その遺 伝子とのペアを大量に用意し、表現型に対応する遺伝子を

1	大阪公立大学 大学院情報学研究科
	Graduate School of Informatics, Osaka Metropolitan Univer
	sity
2	大阪府立大学 生命環境学域

- College of Life, Environment, and Advanced Sciences, Osaka Prefecture University
- ³ 大阪公立大学 大学院農学研究科 Graduate School of Agriculture, Osaka Metropolitan Unviversity
- $^{a)}$ sp25236e@st.omu.ac.jp
- ^{b)} yuzuko@omu.ac.jp
- ^{c)} matsu20@omu.ac.jp
- ^{d)} masa.i@omu.ac.jp



図 1: 根の輪郭

データ解析により求める.他の手法は,交配により遺伝子 の特定を行うため,GWAS は効率的に表現型と遺伝子型の 関連を解析できる.一方で,GWAS を使用するには表現型 の定量化が不可欠である.これまで,総根長や根の表面積 などの表現型を用いて,GWAS を使用した研究が進められ てきた [3],[4].しかし,根の形状を定量化する研究はほと んど行われていない.

根の形状を定量化する既存研究として, Falk らの研究 [5] がある.この研究では,観察が容易な透明な袋の中でダイ ズを栽培し,袋の中で成長したダイズ根の形状を定量化し ている.しかし,ダイズは一般的に圃場で栽培され,屋内 と圃場の栽培では根の成長パターンが異なる [6].そのた め,この既存研究をそのまま適用しても圃場で栽培された ダイズ根の形状を定量化できない.

以上の背景から,本研究では圃場で栽培されたダイズ根 の形状を GWAS で解析可能となるよう,定量化すること を目的とする.定量化の対象は,図1の白い線で示すよう な,根の領域を覆う輪郭とする.はじめに述べたように, IPSJ SIG Technical Report

根の概形だけでも根の栄養吸収の特徴を表現できることか ら、輪郭を用いることで、形状の特徴を適切に表現できる と考えれられる [7]. この輪郭に対して、楕円フーリエ解析 を施し、得られた係数を用いて形状を定量化する.

実験では圃場で栽培されたダイズ根 562 個体を撮影し, 提案手法を適用して,根の形状を定量化した.主成分分析 の結果から,ダイズ根の根の広がりに関連する特徴量を抽 出することができた.さらに,GWAS の結果,本研究で 得た根の形状を表す特徴量と関連のある Single Nucleotide Polymorphism (SNP)が1つ検出された.

2. 関連研究

植物の根の表現型の研究は,栽培環境や,対象とする表 現型の種類によって,多岐にわたる研究が行われている. 本章では,根の表現型の研究で用いられる栽培環境および 表現型について述べる.

2.1 根の表現型に関する研究

根の表現型の研究は、屋内の栽培と圃場の栽培の、2つ の栽培環境に分けられる [8]. 屋内の栽培では、根の観察 が常時可能な、透明のガラスケースやポーチで栽培する方 法 [3], [4], [9] が採用されている. このような屋内の栽培方 法の利点として, 天候の影響を受けにくく, 安定した環境 下で実験データを取得しやすいことが挙げられる.また, 根の状態が常時観察できるため、数時間ごとの定期的な根 の成長の追跡が可能である.しかし、このような屋内の栽 培では、根の周囲がガラスケースなどで囲われているため、 根の成長がケース内に制限される. また、観察は根を掘り 起こさず、ガラスケースに土を入れた状態で行う. そのた め、一部の根が土に隠れた状態で観察することになり、根 の全体の形状を捉えることが困難である [6]. 対して, 圃場 での研究は、栽培が天候に左右されるため、栽培環境が不 安定なことや、根を掘り起こす手間がかかるなどの欠点が 存在する.しかし、圃場は根の成長に制限が少なく、実際 の農業現場に近い環境での栽培が可能である[6].そこで、 本研究では、圃場と屋内の栽培で得られる根の形状が異な るため、圃場で栽培されたダイズ根を対象とする.

植物の根の表現型の研究では,総根長や総根面積といっ た表現型が頻繁に用いられる [10], [11]. これらの表現型 は根の特性を表しているものの,根に関する統計的なデー タであり,2次元形状の情報を失っている.根の形状に関 する指標として,主根と側根の分岐角度を測定している研 究 [12] も存在する.この方法では根の部分的な形状の特徴 を捉えられるものの,根全体の形状を包括的に表現するこ とは難しい.根の形状も植物の特性を表す重要な要素であ ることが指摘されているため [1],本研究では根の表現型と して形状に着目する.

2.2 形状計測の研究

植物に限らず生物の形状を表現するために,様々な手 法が提案されている [13].動物を対象とした研究では,ラ ンドマークを用いた形状解析が行われている [14]. ランド マークとは,対象の特徴的な部位を示す点であり,研究者 が事前に定めるものである.しかし,植物の根は何度も分 枝を繰り返している形状をしており,一意に決まるランド マークを定義することができない.そのため,ランドマー クを用いた手法の適用は難しい [15].

根は、枝分かれが繰り返される構造をしており、一部を 取り出して拡大すると、全体とよく似た形が現れるフラク タル構造を持っている.このことから、フラクタル構造を 持つ図形の、形状の複雑さを定量化する指標であるフラク タル次元を用いて根の形状を表現する手法が提案されてい る [16], [17].しかし、フラクタル次元は、根の形状で養分 の吸収に大きく関わる根の広がりを表現できない.

別の形状の解析方法として,計測対象の輪郭の閉曲面を フーリエ級数展開する楕円フーリエ解析を適用した研究が 報告されている [18], [19].楕円フーリエ解析は,対象の輪 郭情報をフーリエ級数展開し,そのフーリエ係数を用いて 定量的に表現することから,本研究の根形状の定量化に適 している.そのため,本研究では,形状の解析手法として 楕円フーリエ解析を採用する.

根形状の概形を扱った先行研究として,屋内で栽培され た発芽から間もないダイズ根に対する形状定量化の研究が ある [5]. この研究では,発芽から最長 12 日間の根に対し て楕円フーリエ解析を適用して,その結果をクラスタリン グしている.しかし,この研究は屋内で発芽試験紙を用い て栽培されたダイズを対象としており,実際の農業現場の 栽培環境とは異なる.そこで,本研究では,実際の栽培環 境である圃場で約2ヶ月間,露地栽培されたダイズ根に対 して形状解析を行う.

3. 提案手法

本研究では、図 2 の 4 つの手順に従い、ダイズ根の形 状を定量化する.まず、Step 1 の画像補正では、撮影し た画像のカメラレンズによる歪みと、撮影時のカメラの傾 きによる歪みを補正する.次に、Step 2 のセグメンテー ションでは、画像から根の領域を切り出す.このセグメン テーションには、既存の U-Net を用いた根のセグメンテー ションモデル [20] を、本研究の根の画像を用いてファイン チューニングしたものを使用する.続いて、Step 3 の輪郭 抽出では、セグメンテーションで検出した根領域から輪郭 を [21] の手法を用いて抽出し、最も大きいものを根の輪郭 として検出する.最後に、Step 4 の形状の定量化では、楕 円フーリエ解析 [22] を用いて根を楕円フーリエ係数で表現 する.その後、楕円フーリエ係数に対して主成分分析を適 用し、得られた主成分をもとに根の形状を定量化する.以

IPSJ SIG Technical Report



図 2: 提案システムの手順

降, それぞれの Step について詳細に述べる.

3.1 画像補正

画像の撮影は、土から掘り出した根を平面に置き、それ を固定したカメラで撮影する.撮影した画像には、レンズ による歪みと、カメラと撮影対象面が並行でないことから 発生する傾きの歪みが含まれる.これらの歪みにより根の 輪郭が変形するのを防ぐため、Step 1 では画像の補正を行 う.レンズによる歪み補正では、カメラキャリブレーショ ンでカメラの内部パラメータと歪み係数を求め、半径方向 歪みと円周方向歪みの2種類の歪みを補正する.傾きによ る歪み補正では、撮影時に画像内に円形のろ紙を配置し、 撮影した画像上のろ紙を円形にするアフィン変換を施すこ とで補正する.以降、レンズ歪み補正とカメラの傾きによ る歪みの補正について詳細に説明する.

3.1.1 レンズ歪み補正

通常,カメラで撮影した画像には、レンズを用いること により歪みが生じる.これらの歪みは画像上の根の形状を 変化させ,根の概形を正確に抽出することができないため, これらの歪みの補正が必要である.本研究では,歪み補正 として,Brown[23]やBrownら[24]が導出したモデルを用 いる.このモデルでは,レンズ歪みの主な原因である,半 径方向歪みと円周方向歪みの2種類の歪みを扱っている.

半径方向歪みは、レンズの曲面形状によって、光線の投 射位置が受像素子上で移動することで発生する歪みであ る.一般に、単レンズでは光軸から離れるほど投射位置の 歪みによる移動量は大きくなり、その方向は光軸を中心に 等方的である.画像平面上の点 (x,y)における半径方向歪 みによる移動量 Δx_r , Δy_r は、 $\bar{x} = x - c_x$, $\bar{y} = y - c_y$ を画 像の中心座標 c_x , c_y を基準とした座標、 $r^2 = x^2 + y^2$ を光 軸から点 (x,y) までの距離、 k_1, k_2, k_3, \cdots を半径方向歪み 係数とすると、高次多項式の偶数乗項を用いた式 (1), (2) で表わされる.

$$\Delta x_r = \left(k_1 r^2 + k_2 r^4 + k_3 r^6 + \cdots\right) \bar{x} \tag{1}$$

$$\Delta y_r = \left(k_1 r^2 + k_2 r^4 + k_3 r^6 + \cdots\right) \bar{y} \tag{2}$$

円周方向歪みは、レンズと受像素子が完全に平行でない ことによって発生する歪みである.この歪みによる移動 量 $\Delta x_t, \Delta y_t$ は、 p_1, p_2, \ldots を円周方向歪み係数とすると、 式 (3), (4) で表される.

$$\Delta x_t = \left\{ 2p_1 \bar{x} \bar{y} + p_2 \left(r^2 + 2\bar{x}^2 \right) \right\} \left(1 + p_3 r^2 + \cdots \right) \quad (3)$$
$$\Delta y_t = \left\{ p_1 \left(r^2 + 2\bar{y}^2 \right) + 2p_2 \bar{x} \bar{y} \right\} \left(1 + p_3 r^2 + \cdots \right) \quad (4)$$

これら2種類の歪みの補正は, x', y'を補正後の座標, f_x, f_y を焦点距離, $\bar{x} = \frac{x-c_x}{f_x}, \bar{y} = \frac{y-c_y}{f_y}$ を歪みを含む点とすると, 式 (5), (6) で表される.

$$x' = \frac{\bar{x} - \left\{2p_1\bar{x}\bar{y} + p_2\left(r^2 + 2\bar{x}^2\right)\right\} \left(1 + p_3r^2 + \cdots\right)}{k_1r^2 + k_2r^4 + k_3r^6 + \cdots}$$
(5)
$$y' = \frac{\bar{y} - \left\{p_1\left(r^2 + 2\bar{y}^2\right) + 2p_2\bar{x}\bar{y}\right\} \left(1 + p_3r^2 + \cdots\right)}{k_1r^2 + k_2r^4 + k_3r^6 + \cdots}$$
(6)

本研究では、計算量を削減するため、半径方向歪み係数を k_1, k_2, k_3 、円周方向歪み係数を p_1, p_2 のみを考慮し、これら以外の係数を0とした歪みモデルを利用する.

カメラの内部パラメータ (c_x, c_y, f_x, f_y) と歪み係数 $(k_1, k_2, k_3, p_1, p_2)$ は、Zhang の手法 [25] を基にしたカメ ラキャリブレーションによって求める. この手法では、 チェッカーボードなどの平面物体上のパターンを様々な方 向に向けて撮影し、画像からチェッカーボードのコーナー を検出する. その後、検出したコーナーの再投影誤差を最 小化し、カメラの内部パラメータ、外部パラメータ、歪み 係数を推定する. N 個の特徴点を持つ平面物体を撮影し た画像が K 枚存在するとき、A をカメラ内部パラメータ 行列、D をレンズ歪み係数 $(k_1, k_2, k_3, p_1, p_2)$, $R_i, T_i \in i$ 番目の画像における外部パラメータ、 m_{ij} を画像中での特 徴点の 2 次元の座標データ、 M_{ij} を平面物体上での 3 次元 の座標データ、関数 m を与えられたパラメータに基づい て受像素子上の点に点 M_{ij} を投影する再投影関数とする. すると、再投影誤差 error は、式 (7) で表される.

error =
$$\sum_{i}^{K} \sum_{j}^{N} ||m_{ij} - \breve{m}(A, D, R_i, T_i, M_{ij})||^2$$
 (7)

再投影誤差を最小化するための反復法は Levenberg-Marquardt 法 [26] が用いられる.

3.1.2 傾き補正

画像の撮影時にカメラと対象物を完全に平行にするのは 難しく,傾きが生じる.その結果,対象物が意図しない形 に変形してしまう.このように変形した形状を,そのまま 定量化することを防ぐため,傾き補正をする.

本研究では、図 3(a) のように、傾き補正を行うために、 撮影画像内に円形のろ紙を入れ、このろ紙を真円とみなす. カメラの撮像面と根を置いた平面は平行になっていないた め、ろ紙は歪み、楕円として撮影される.そこで、画像か らろ紙を検出し、ろ紙の形状を楕円から真円にする画像変 換をすることで、傾きを補正する.

本研究で行う傾き補正の手順を図3に示す.まず,図3(a)



に示す画像に対して楕円フィッティング [27] を用いて,楕 円の中心点,長軸の長さ,短軸の長さ,傾きの角度を求め る.次に,図 3(b)のように,回転させた時に元の画像が 画像領域内に収まるよう,出力画像の大きさを縦横 2 倍に 拡大し,楕円の中心点を基準に,楕円の長軸と短軸の方向 が x,y 軸に一致するよう回転させる.そして,図 3(c)の ように,長軸の長さが短軸に等しくなるよう,画像全体を アフィン変換する.最後に,図 3(d)のように,楕円の中心 を基準に,元の傾きに戻るように画像を回転させてから, 出力画像を元の大きさに戻す.

3.2 セグメンテーション

Step 2 のセグメンテーションでは, Smith らの U-Net を 用いたセグメンテーションモデル [20] をファインチューニ ングしたモデルを使用する. Smith らのセグメンテーショ ンモデルは, U-Net の入力を RGB 画像に変更した点を除 いては, 元の U-Net[28] と同じ構造を持つ.

3.2.1 U-Net の概要

U-Net とは, Ronneberger らによって生物医学の画像セグ メンテーションのために開発された Convolutional Neural Network (CNN) である. U-Net は,構造がアルファベット の「U」の形に似ていることからこの名がついている [28].

U-Net はエンコーダとデコーダの2つの部分から構成さ れ、全結合層を用いずに畳み込み層のみで設計されている. エンコーダでは入力画像を縮小しながら特徴を抽出し、デ コーダではエンコーダで縮小された画像を拡大しながら元 の解像度に近づけている.しかし、このような拡大処理で は空間情報の復元が不完全になり、精度の低下を招くこと がある.この問題を解決するために、U-Net ではエンコー ダとデコーダの対応する層同士をスキップ接続している. このスキップ接続によって、エンコーダで得られた特徴 マップをデコーダで再利用できるため、空間情報などの特 徴を保持したまま画像を復元できる.

損失関数には重み付き交差エントロピー誤差を使用す る.重み付き交差エントロピー誤差とは、オブジェクトの 境界部分に高い重みを与えることで、境界領域の誤差を重 視した交差エントロピー誤差である.最適化には、確率的 勾配法を用いる.

3.2.2 データ拡張

本研究で得られる根画像は,栽培や撮影の労力から数百 枚程度に限られる.そこで,少ないデータでの学習を可能



図 4: 輪郭抽出の手順

にし、過学習を防ぐためにデータ拡張を用いる.

本研究では、U-Netと同様に、データ拡張に画像全体に 局所的な歪みを加える変換である弾性変換 (Elastic Distortions)[29] を用いる.手順としては、まず、全ての座標の移 動範囲を決め、決めた範囲内で各ピクセルごとに移動量を 一様乱数で決める.次に、その値を 2 次元配列に格納し、 ガウシアンフィルタを適用して平滑化する.このガウシア ンフィルタの標準偏差 σ は、弾性変換の平滑性を表すパラ メータとなる.そして、平滑化した配列に変形のスケール 係数 α をかけ、各ピクセルの座標を移動させる.最後に、 移動後の座標が整数でない場合、ピクセルの位置を滑らか に補正するためにバイリニア補完を行う.

弾性変換に加え,明るさ,コントラスト,彩度,色相を 変更し,ガウスノイズやごま塩ノイズの適用,画像の左右 反転をデータ拡張操作として利用する.

3.3 輪郭抽出

Step 3 の輪郭抽出では,手法として Suzuki らのアルゴ リズム [21] を用いる.このアルゴリズムは二値画像に対 し,ラスタスキャンを適用し,輪郭を抽出する手法である.

手法の手順を図4で示す. 白く塗りつぶされたマスは前 景,黒く塗りつぶされたマスは背景を表している. 黄色で 塗りつぶされたマスは検出された輪郭の画素を表し、オレ ンジ色で塗りつぶされたマスは次に検出される輪郭の画素 を示す.青の矢印において,矢の始点が現在の画素,矢の 終端が追跡される画素である.図4(a)の青の矢印で示す ように、 ラスタスキャンにより、 画像の各行を左から右へ 走査し、下の行へと順に移動する. そして、図 4(b)の黄色 で示すように、ラスタスキャンの過程で前景が検出された 場合,その画素を起点として輪郭の追跡を開始する.輪郭 の追跡は、対象画素の周囲8近傍を時計回りに探索し、最 初に発見された前景の画素へ移動する. この操作を繰り返 すことで、図 4(c) のように、輪郭を連続的にたどる. この 輪郭の追跡は、端点に到達するか、初期位置に戻るまで継 続される.輪郭の追跡が終わると,図 4(d) のように,ラ スタスキャンを再開し、まだどの輪郭にも属さない画素を 検出した場合,新たな輪郭として追跡処理を行う.この手 順を繰り返すことで、画像内の輪郭を抽出する.

3.4 形状の定量化

Step 4の形状の定量化では、楕円フーリエ解析と主成分

IPSJ SIG Technical Report

分析を用いる.楕円フーリエ解析では,輪郭を2つの正弦 波の和で表現し,楕円フーリエ級数によって輪郭の形状を 表現する.主成分分析では,楕円フーリエ解析によって得 られたフーリエ係数の次元を削減し,データの主要な特徴 を抽出する.この特徴をもとに,根の形状を定量化する.

3.4.1 楕円フーリエ解析

本研究で根の形状の定量化手法として用いる楕円フーリ エ解析 [22] は、輪郭形状を周期関数とみなし、フーリエ級 数展開する手法である.楕円フーリエ解析では、図 5 で示 すように、輪郭上のある点 *S* から出発し、輪郭上を一定の 速度で周回する点 *P* を考える.点 *P* の時刻 *t* における座 標を (x(t), y(t)) とすると、点 *P* は輪郭上を 1 周するごと に出発点 *S* に戻るため、x(t), y(t) は周回時間 *T* を周期と する周期関数となる.このような周期関数はフーリエ級数 により記述可能であり、 A_0 と C_0 をそれぞれ x 軸と y 軸 方向の定数項、 A_n, B_n, C_n, D_n を形状の特徴を表すフーリ エ係数、*N* をフーリエ級数の次数とすると、x(t), y(t) は それぞれ式 (8)、(9) で表される.

$$x(t) = A_0 + \sum_{n=1}^{N} \left(A_n \cos \frac{2\pi nt}{T} + B_n \sin \frac{2\pi nt}{T} \right)$$
(8)
$$y(t) = C_0 + \sum_{n=1}^{N} \left(C_n \cos \frac{2\pi nt}{T} + D_n \sin \frac{2\pi nt}{T} \right)$$
(9)

この式から、 A_n, B_n, C_n, D_n の 4 種類のフーリエ係数を N個ずつ得られ、合計 4N 個のフーリエ係数が根の形状を表 す特徴量となる.フーリエ級数の次数 N を増やすと、輪 郭の近似精度が向上する.

本研究では、楕円フーリエ解析を行う前に、輪郭に対し て位置調整とスケーリングを行う. 位置調整では、根のセ グメンテーション結果の重心が同じになるように、根を平 行移動させる. スケーリングでは、*x*,*y* を変換前の根の輪 郭の座標列,*x'*,*y'* を変換後の座標列,*A* を根の輪郭の面 積として,式(10)で根の輪郭の大きさを揃える.

$$\boldsymbol{x}' = \frac{\boldsymbol{x}}{\sqrt{A}}, \quad \boldsymbol{y}' = \frac{\boldsymbol{y}}{\sqrt{A}}$$
 (10)

ここで,輪郭の面積を平方根を取っているのは,面積の次 元を座標と同じ次元に揃えるためである.この前処理によ り,楕円フーリエ解析で得られるフーリエ係数は,面積の 大きさや位置に依存せず,形状の特徴を抽出できる.

3.4.2 主成分分析

楕円フーリエ解析によって得られるフーリエ係数は, N の数を増やしてやれば詳細な形状を記述可能な一方で,係 数が増えることでデータ解析が煩雑になる.そこで本研究 では,楕円フーリエ解析を用いて得られたフーリエ係数の 次元を削減し,根の形状の主要な特徴量を抽出するために, 主成分分析 [30] を適用する.

主成分分析では,楕円フーリエ解析で得られたフーリ



図 5: 楕円フーリエ解析の輪郭の表現方法. 横軸が x 座標, 縦軸が y 座標を表す. 点 S から出発し, 黄色の矢印の方向 に点 P が輪郭上を周回する.

エ級数の係数 (A₀, B₀, C₀, D₀,...) を得られたデータとし, 共分散行列を求め,その固有値・固有ベクトルを計算する. 得られた固有ベクトルは主成分として扱い,寄与率が高い 順に並び替える.この寄与率は共分散行列の固有値に基づ いて算出される.寄与率の高い主成分を用いることで,元 のデータの情報をできる限り保持しつつ,次元削減ができ る.この手法を,フーリエ係数に適用することで,楕円フー リエ解析で得られた根の形状データを低次元に変換し,形 状の特徴を抽出する.

4. 実験

提案手法の検証のため, 圃場で栽培されたダイズ根の画 像を使用し,実験を行った.本章では実験データ,実験環 境,実験結果を述べ,結果について考察する.

4.1 実験データ

実験では、大阪公立大学農学部附属教育研究フィールド の灰色低地土の圃場で, 2023年9月11日から11月28日 まで栽培されたダイズを用いた. 圃場には、リン無施肥区 とリン施肥区の2つの処理区を設けた.ダイズは、農研機 構が配布している日本のダイズコアコレクション*196品 種のうち,種子増殖が完了している 94 品種を使用し,各 品種3個体ずつ栽培した.予定では564個体のダイズを 撮影する予定であったが、2個体が枯死したため、最終的 には 562 個体を撮影した.撮影には、三脚に固定したカメ ラ CANON EOS RP を使用した.撮影環境は、図 6(a) に 示すように、テントの骨組みに黒のビニールシートを取り 付け,光の入射を最小限に抑えた.三脚はテント内に設置 し、ノートパソコンでカメラを遠隔操作して撮影した. 傾 き補正および大きさの比較のため、直径 9cm の円形のろ 紙を画像の左上に置き,根とともに撮影した.画像の解像 度は 4160 × 4160 ピクセル, 画像の形式は RGB カラー画 像である.撮影した画像の例を図 6(b) に示す. 歪み補正 には、カメラキャリブレーションを実施するために、1辺

^{*1} 日本のダイズコアコレクション, https://www.gene.affrc.go. jp/databases-core_collections_jg.php

IPSJ SIG Technical Report



(a) 撮影環境

(b) 撮影した画像図 6: 実験データ

が 22mm の 11×8 のマス目を持つチェッカーボードを使 用した.根の撮影と同様の環境で、チェッカーボードの角 度を変化させて画像を撮影した.合計で 10 枚の画像を撮 影し、実験に使用した.

4.2 実験環境

本研究で用いた計算機は、画像補正、輪郭抽出、形状の定 量化では、MacBook Pro 16-inch 2019 (8 コア、32GB メモ リ、1TB ハードディスク)を使用した. セグメンテーション では GPU NVIDIA RTX A6000AMD と CPU FPYC 74F3 を搭載したサーバーを使用した. 画像補正と輪郭抽出には Python 3.9.6 を使用し、画像処理ライブラリの OpenCV 4.10.0*2を用いた. セグメンテーションには Python 3.10.14 を使用した. 楕円フーリエ解析と主成分分析は R 4.4.1 を 用い、形状解析ライブラリの Momocs 1.4.1*3を用いた. 4.2.1 セグメンテーション

セグメンテーションモデルのファインチューニングや, 精度評価に利用する画像として,4.1 で説明した562 枚の 画像から,18 枚を選定した.学習画像での系統や栽培環境 の偏りを防ぐため,17 系統のダイズ根の画像から選定し, 栽培環境もリン無施肥区とリン施肥区の2つの環境が均等 になるように選定した.根領域の正解データは,Adobe社 製 PhotoShop 25.9 の閾値処理を用いて,根と背景を大ま かに分類した後,手作業でピクセル単位の詳細な根領域の 正解ラベル付けを実施した.作成した18 枚の正解を付与 したデータは,ファインチューニング用の訓練データ14 枚,検証データ2枚,テストデータ2枚に振り分けて使用 した.セグメンテーションモデルの学習パラメータとデー タ拡張のパラメータは,Smithらのセグメンテーションモ デル[20] の学習時と同じ条件に設定した.

根のセグメンテーションモデルの評価指標として F1 ス コアを用いた. F1 スコアは,陽性と陰性のピクセルの割合 に偏りがある場合でも,精度と再現率のバランスを取るの

© 1959 Information Processing Society of Japan

に有効な指標である. F1 スコアは, Precision と Recall を 用いて計算される. Precision と Recall は, True Positive (TP)を正解とモデルの予測値の両方が陽性であるピクセル 数, True Negative (TN)を正解とモデルの予測値の両方が 陰性であるピクセル数, False Positive (FP)を正解が陰性 でモデルの予想値が陽性であるピクセル数, False Negative (FN)を正解が陽性でモデルの予想値が陰性であるピクセ ル数とすると, それぞれ式 (11), (12) で表される.

$$Precision = \frac{TP}{TP + FP}$$
(11)

$$Recall = \frac{TP}{TP + FN}$$
(12)

そして F1 スコアは, これらの指標の調和平均を取った, 式 (13) で表される.

$$F1 \ \mathcal{A} \ \exists \ \mathcal{P} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$
(13)

4.2.2 輪郭抽出

輪郭抽出には、3.3 を用いて根をセグメンテーションし た画像を使用した.根領域のセグメンテーション時に誤検 出される背景の布の折り目や、キャリブレーション用に置 いたろ紙が誤検出されている.これらの誤検出結果は誤っ た輪郭抽出を招く可能性がある.そこで、画像の右上隅か ら 3500 × 400 ピクセルと 1300 × 1000 ピクセルの 2 つの 領域の検出結果を全て背景とした.また、輪郭抽出の際、 ろ紙の輪郭や布の折り目以外の部位での根の誤検出の影響 から、複数の輪郭が抽出される.誤検出されたものの輪郭 は小さく、根の輪郭は大きいことから、検出された輪郭の 中で面積が最も大きいものを抽出結果とした.評価方法は セグメンテーション結果の画像と輪郭抽出結果の画像を目 視で比較する定性評価を行った.

4.2.3 形状の定量化

楕円フーリエ解析におけるフーリエ級数の次数 N は,元 の図形を十分に表現できない事態を避けるため,フーリエ 級数から復元した輪郭を目視で確認し,輪郭の変化が現れ なくなった N=100 に設定した.

4.3 結果と考察

4.3.1 画像補正

画像補正の結果を図7に示す. 歪み補正で求めた再投影 誤差は0.04 ピクセルであった. 輪郭で根の形状を扱う本 研究において、1 ピクセル以下の歪みによる誤差は無視で きるほど小さな誤差であるため、歪み補正の精度は十分で あると評価できる. 傾き補正では、変換前の画像では楕円 上になっていたろ紙を、直径が629 ピクセルの円に統一し て変換することで、傾きを補正した.

4.3.2 セグメンテーション

根のセグメンテーション結果は, Precision は 0.636, Recall は 0.831, F1 スコアは 0.721 であった. 根のセグメン

^{*2} OpenCV:Open Source Computer Vision Library, https:// opencv.org/

^{*3} Momocs: Open Source 2D Morphometrics, https://momx.github.io/Momocs/index.html

IPSJ SIG Technical Report



(a) 入力画像 (b) 歪み補正の結果 (c) 傾き補正の結果 図 7: 画像補正の結果





図 8: 根のセグメンテーション結果

セグメンテーションの結果

入力画像



図 9: 輪郭抽出の結果

テーション結果を図8に示す.この結果から, precision が 低い原因は、ろ紙の輪郭や背景に使った布の折り目の誤検 出が影響していることが考えられる.本研究では、根の輪 郭のみを形状の解析に使うため、図8に示すような誤検出 は根の概形に影響せず,許容できる. Recall は高いことか ら、検出できなかった根の領域は少ないことが分かる. そ のため、根の領域は概ね抽出できている.

4.3.3 輪郭抽出

輪郭抽出の結果を図9に示す.全ての根において、根の 全体が欠けるような大きな欠損は見られなかった. そのた め、輪郭抽出ではおおよその根の形状を表現することがで きている.

4.3.4 形状の定量化

楕円フーリエ解析で得られたフーリエ係数に対し, 主成 分分析を適用した. 第一主成分と第二主成分の寄与率はそ れぞれ 0.280%, 0.248%で, 第一主成分と第二主成分で全体 の成分の 50%を占めていることがわかる. その際の第一主 成分と第二主成分をそれぞれ x, y 軸として, 実験で用い た根画像の輪郭データをプロットした散布図を図 10 に示 す.図 10 の背景に並んで表示されている輪郭はグラフ上 の各位置に対応する形状を表している.図10より,第一 主成分を見ると,値が大きい時は根の分岐がなく,細長く



図 10: 第一主成分と第二主成分を軸とする主成分分析結果



図 11: 第一主成分に対するゲノムワイド関連解析の結果

伸びた形状を示し、値が小さい時は根の分岐が発生し、広 がった形状を示している.そして、第二主成分を見ると、 値が大きい時は根が右に偏った形状になり、値が小さい時 は根が左に偏った形状になっていた. ここで, 第二主成分 の特徴は、撮影時の根とカメラの位置による影響を受ける ため、根の形状そのものの特徴とは言い難い、そこで、本 研究では第一主成分のみが根の形状を表現していると考 え, GWAS に使用した.

第一主成分についてゲノムワイド関連解析を行い, 遺伝 子との関連を調べた結果を図 11 に示す. このグラフの各 点は遺伝子の変異である. 横軸は遺伝子の変異が起きてい る部位である Single Nucleotide Polymorphism (SNP)の 染色体上での位置を示しており,縦軸は SNP と第一主成分 との関連の強さを表している p 値である.そして、点線は、 p 値の統計的有優位水準であり、今回の閾値は 2.35 × 10⁻⁶ とした.. この値を超えた SNP は第一主成分と関連がある といえる.図 11 より、P 値を超える SNP が一つ確認で き、本研究で得た形状を表す特徴量と関連のある SNP が 一つ検出された. この結果から、本研究では圃場で栽培さ れたダイズ根の形状を定量化し、根の形状と遺伝子型との 関連のある SNP を検出できた.

5. 結論

任意の根形状を持つ植物の育種を効率的に行う方法とし て,遺伝子操作による改良が挙げられる.遺伝子操作によ る育種では、根の形状と遺伝子型との対応を明らかにす る必要がある.根の形状と遺伝子型との対応を取るには、

IPSJ SIG Technical Report

GWAS が最も手軽で容易な手法である.しかし,その適用 には根の形状を定量化する必要がある.根の形状に関する 従来研究では,屋内栽培された観察しやすい根を用いたも のが多く,これらの根の形状は圃場で栽培されたものと大 幅に異なる.根形状は栽培環境により形状が大きく変化す る.そのため,圃場で栽培されるダイズの根形状を解析す る必要がある.そこで,本研究では,根の形状と遺伝子型 との関連を解析するために,圃場で栽培されたダイズ根に 対し,楕円フーリエ解析を用いた形状の定量化を行った.

圃場で栽培された根の画像を用いた実験で,根の伸び方 と広がりに関する形状の特徴をもとに,根形状を定量化し た.さらに,ゲノムワイド関連解析により,本研究で得た 根の形状を表す特徴量と関連のある SNP が検出された. これらの結果から,本研究の手法により,圃場で栽培され たダイズ根の形状を定量化し,根の形状と遺伝子型との関 連の解析に貢献できることが示された.今後は,他の品種 のダイズ根の画像にも本手法を適用することを目指す.

参考文献

- Uga, Y.: Challenges to design-oriented breeding of root system architecture adapted to climate change, *Breed. Sci.*, Vol. 71, No. 1, pp. 3–12 (2021).
- [2] Ozaki, K. et al.: Functional SNPs in the lymphotoxinalpha gene that are associated with susceptibility to myocardial infarction, *Nat. Genet.*, Vol. 32, pp. 650–654 (2002).
- [3] Zhao, H. et al.: RhizoPot platform: A high-throughput in situ root phenotyping platform with integrated hardware and software, *Front. Plant Sci.*, Vol. 13, No. 1004904, pp. 1–13 (2022).
- [4] Chandnani, R. et al.: Application of an Improved 2-Dimensional High-Throughput Soybean Root Phenotyping Platform to Identify Novel Genetic Variants Regulating Root Architecture Traits, *Plant Phenomics*, Vol. 5, No. 0097, pp. 1–18 (2023).
- [5] Falk, K. G. et al.: Computer vision and machine learning enabled soybean root phenotyping pipeline, *Plant Meth*ods, Vol. 16, No. 5, pp. 1–19 (2020).
- Liu, S. et al.: DIRT/3D: 3D root phenotyping for fieldgrown maize (Zea mays), *Plant Physiol.*, Vol. 187, No. 2, pp. 739–757 (2021).
- [7] De Winter, J. et al.: Contour-based object identification and segmentation: Stimuli, norms and data, and software tools, *Beh. Res. Meth. Instr. Comp.*, Vol. 36, No. 4, pp. 604–624 (2004).
- [8] Paez-Garcia, A. et al.: Root Traits and Phenotyping Strategies for Plant Improvement, *Plants*, Vol. 4, No. 2, pp. 334–355 (2015).
- [9] Khoroshevsky, F. et al.: Automatic Root Length Estimation from Images Acquired In Situ without Segmentation, *Plant Phenomics*, Vol. 6, No. 0132, pp. 1–15 (2024).
- [10] Narisetti, N. et al.: Fully-automated root image analysis (faRIA), Sci. Rep., Vol. 11, No. 16047, pp. 1–15 (2021).
- [11] Seethepalli, A. et al.: RhizoVision Explorer: open-source software for root image analysis and measurement standardization, AoB PLANTS, Vol. 13, No. 6, plab056, pp. 1–15 (2021).
- [12] Coronado-Aleans, V. et al.: High-throughput Phenotyp-

ing of Maize Roots Using Digital Image Analysis, *Ciencia y Tecnología Agropecuaria*, Vol. 25, No. 1, pp. 1–16 (2024).

- [13] Mitteroecker, P. et al.: Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness, *AJBA*, Vol. 178, No. S74, pp. 181–210 (2022).
- [14] Dwivedi, A. K. et al.: Role of Morphometrics in Fish Diversity Assessment: Status, Challenges and Future Prospects, *Natl. Acad. Sci. Lett.*, Vol. 47, No. 2, pp. 123–126 (2024).
- [15] Yasrab, R. et al.: RootNav 2.0: Deep learning for automatic navigation of complex plant root architectures, *GigaScience*, Vol. 8, No. 11, giz123, pp. 1–16 (2019).
- [16] Peeples, J. et al.: Spatial and Texture Analysis of Root System distribution with Earth mover's Distance (STARSEED), *Plant Methods*, Vol. 19, No. 2, pp. 1–15 (2023).
- [17] Belzile, F. et al.: Comparing Results from 2-D and 3-D Phenotyping Systems for Soybean Root System Architecture: A 'Comparison of Apples and Oranges' ?, *Plants*, Vol. 13, No. 23, 3369, pp. 1–8 (2024).
- [18] García-Bustos, M. et al.: New Methods for Old Questions: The Use of Elliptic Fourier Analysis for the Formal Study of Palaeolithic Art, J. Archaeol. Method and Theory, Vol. 31, No. 4, pp. 1843–1872 (2024).
- [19] Jeanty, A. et al.: Identification of archaeological barley grains using geometric morphometrics and experimental charring, J. Archaeol. Sci., Vol. 162, No. 105924, pp. 1–25 (2024).
- [20] Smith, A. G. et al.: Segmentation of roots in soil with U-Net, *Plant Methods*, Vol. 16, No. 13, pp. 1–15 (2020).
- [21] Suzuki, S. et al.: Topological structural analysis of digitized binary images by border following, *Comput. Gr. Image Process.*, Vol. 30, No. 1, pp. 32–46 (1985).
- [22] Kuhl, F. P. et al.: Elliptic Fourier features of a closed contour, *Computer Graphics and Image Processing*, Vol. 18, No. 3, pp. 236–258 (1982).
- [23] Duane, C. B.: Close-Range Camera Calibration, Photogramm. Eng., Vol. 37, No. 8, pp. 855–866 (1971).
- [24] Fryer, J. G. et al.: Lens Distortion for Close-Range Photogrammetry, *Photogramm. Eng. Remote Sens.*, Vol. 52, pp. 51–58 (1986).
- [25] Zhang, Z.: A flexible new technique for camera calibration, *IEEE Trans. on PAMI*, Vol. 22, No. 11, pp. 1330– 1334 (2000).
- [26] Ananth, R.: The Levenberg-Marquardt Algorithm, *Tu-toral on LM algorithm*, Vol. 11, No. 1, pp. 101–110 (2004).
- [27] Fitzgibbon, A. et al.: Direct least square fitting of ellipses, *IEEE Trans. on PAMI*, Vol. 21, No. 5, pp. 476– 480 (1999).
- [28] Ronneberger, O. et al.: U-Net: Convolutional Networks for Biomedical Image Segmentation, *Proc. of MICCAI*, pp. 234–241 (2015).
- [29] Simard, P. et al.: Best practices for convolutional neural networks applied to visual document analysis, *Proc. of ICDAR*, pp. 958–963 (2003).
- [30] Maćkiewicz, A. et al.: Principal components analysis (PCA), Comput. Geosci., Vol. 19, No. 3, pp. 303–342 (1993).