

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра клеточной биологии и биоинженерии растений**

САВИЦКИЙ

Артём Сергеевич

**РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ
ФЕНОТИПИЧЕСКОГО АНАЛИЗА РАСТЕНИЙ *ARABIDOPSIS
THALIANA* В КУЛЬТУРЕ *IN VITRO***

Магистерская диссертация

специальность 7-06-0511-05 «Биоинформатика»

Научный руководитель:
Демидчик Вадим Викторович
член-корр. НАН Беларуси,
доктор биологических наук, профессор

Допущена к защите

«___» _____ 2024 г.

Зав. кафедрой клеточной биологии и биоинженерии растений

кандидат биологических наук, доцент

_____ О.Г. Яковец

Минск, 2024

СОДЕРЖАНИЕ

Реферат.....	4
Введение.....	7
Глава 1. Литературный обзор.....	9
1.1 Особенности цифрового фенотипирования растений	9
1.2 Проблемы анализа фенотипов растений.....	13
1.2.1 Использование технологии получения изображений в видимом спектре для анализа корней растения.....	15
1.2.2 Использование технологии получения изображений в видимом спектре для анализа надземной части растений.....	16
1.3 Цифровое фенотипирование растений.....	18
1.3.1 Особенности фенотипирования листьев	19
1.3.2 Анализ цифровых изображений корней.....	21
1.3.3 Использование методов фенотипирования для изучения реакции растений на стресс.....	22
Глава 2. Объект и методы исследования	25
2.1 Описание растения <i>Arabidopsis thaliana</i> (L.) Heynh. В качестве модельного объекта для исследования.....	25
2.1.1 Ботаническое описание	25
2.1.2. Генетические особенности <i>Arabidopsis thaliana</i> (L.) Heynh.....	27
2.2 Культивирование <i>Arabidopsis thaliana in vitro</i>	28
2.3 Сбор и анализ изображений.....	29
2.4 Использование библиотеки компьютерного зрения opencv для цифрового фенотипирования	29
2.4.1 Предварительная обработка цифровых изображений.....	31
2.4.2 Методы преобразования цифровых изображений.....	33
2.4.3. Поиск контуров растительных объектов на изображении и их отделение от фона.....	36
2.4.4. Использование метода k-средних для подготовки к отделению объектов от фона	37
2.5. Использование библиотеки машинного обучения “scikit-learn” для цифрового фенотипирования.....	38

2.5.1 Построение модели для предсказания биомассы растений.....	39
2.6 Статистическая обработка данных.....	40
Глава 3 особенности работы программного комплекса для фенотипирования.....	42
3.1 Разработка подходов для определения площадей поверхности органов растений.....	42
3.2 Создание предсказательной модели для оценки биомассы растений....	48
Глава 4. Оценка изменения ростовых параметров <i>Arabidopsis thaliana</i> L. Heynh.....	50
4.1 Анализ видимой зеленой части <i>Arabidopsis thaliana</i> L. Heynh.....	50
4.2 Измерение морфологических параметров корней <i>Arabidopsis thaliana</i> L. Heynh.....	55
4.3 определение изменения морфологических параметров <i>Arabidopsis</i> <i>thaliana</i> L. Heynh на среде с добавлением смеси, генерирующей гидроксил- радикал.....	59
Заключение.....	63

РЕФЕРАТ

Магистерская работа 71 с., 20 рис., 93 источников.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ФЕНОТИПИЧЕСКОГО АНАЛИЗА РАСТЕНИЙ *ARABIDOPSIS THALIANA* В КУЛЬТУРЕ *IN VITRO*

Объект исследования: цифровые фотографии растений *Arabidopsis thaliana* L. Heynh., культура растений *Arabidopsis thaliana* L. Heynh.

Цель исследования: разработка программного обеспечения, позволяющего автоматизировать предобработку и анализ биологических фенотипических данных для растущей культуры *Arabidopsis thaliana* L. Heynh..

Методы исследования: язык программирования Python, методы машинного обучения и построения предсказательных моделей, техника работы с культурой *Arabidopsis thaliana*, получение изображений растений *Arabidopsis thaliana* L. Heynh. в культуре *in vitro*, морфометрические методы определения ростовых параметров, анализ растений с помощью разработанного программного обеспечения.

Результаты: разработан программный комплекс для феномного анализа растений *Arabidopsis thaliana* L. Heynh. на основе библиотек компьютерного зрения и машинного обучения. Разработанный комплекс позволяет измерять площади поверхностей органов растений, а также предсказывать их сырую и сухую массу. С использованием комплекса проанализирована выборка из 35 чашек с культурой *in vitro* молодых растений *Arabidopsis thaliana*, выращенных в контрольных условиях и в условиях окислительного стресса. Измерены значения площадей корней и листьев по данным, извлеченным на основе изображений, получаемых ежедневно на протяжении всего периода культивирования. Эти параметры в дальнейшем использованы для предсказания масс корней и листьев с участием обученной модели, точность которой составила 96,6% по методу R-квадрата. В результате произведены вычисления доли масс корней и листьев по отношению к общей массе растения, соотношение площади корней и листьев к их сухим массам, а также изменение этих параметров в течение периода культивирования. На основе данных феномного анализа показано, что окислительный стресс, вызываемый смесями, генерирующими гидроксильные радикалы, вызывает статистически-достоверное снижение скорости набора биомассы корней и листьев растений *Arabidopsis thaliana* L. Heynh.

ABSTRACT

Master's thesis 71 p., 20 figures, 93 sources.

DEVELOPMENT OF A SOFTWARE PACKAGE FOR PHENOTYPIC ANALYSIS OF ARABIDOPSIS THALIANA PLANTS *IN VITRO* CULTURE

Object of research: Digital photographs of *Arabidopsis thaliana* L. Heynh. plants, *Arabidopsis thaliana* L. Heynh. plant culture.

Purpose of research: Development of software to automate the preprocessing and analysis of biological phenotypic data for a growing culture of *Arabidopsis thaliana* L. Heynh.

Methods of research: Python programming language, methods of machine learning and construction of predictive models, techniques of working with *Arabidopsis thaliana* culture, imaging of *Arabidopsis thaliana* L. Heynh. plants in in vitro culture, morphometric methods of determining growth parameters, analysis of plants using the developed software.

Obtained results: A software package for phenomic analysis of *Arabidopsis thaliana* L. Heynh. plants based on computer vision and machine learning libraries has been developed. The developed complex allows measuring surface areas of plant organs and predicting their wet and dry masses. Using the complex, a sample of 35 in vitro culture cups of young *Arabidopsis thaliana* plants grown under control conditions and under oxidative stress was analyzed. Root and leaf area values were measured from data extracted from images acquired daily throughout the cultivation period. These parameters were further used to predict root and leaf masses with the trained model, which had an R-squared accuracy of 96.6%. As a result, the proportion of root and leaf masses in relation to the total plant mass, the ratio of root and leaf area to their dry masses, and the change in these parameters during the cultivation period were calculated. Based on the data of phenomic analysis, it was shown that oxidative stress induced by mixtures generating hydroxyl radicals causes a statistically significant decrease in the rate of biomass recruitment of roots and leaves of *Arabidopsis thaliana* L. Heynh. plants.

РЭФЕРАТ

Магістарская дысэртацыя 71 с., 20 мал., 93 крыніц.

РАСПРАЦОЎКА ПРАГРАМНАГА КОМПЛЕКСУ ДЛЯ ФЕНАТЫПІЧНАЕ АНАЛІЗУ РАСЛІН *ARABIDOPSIS THALIANA* Ў КУЛЬТУРЫ *IN VITRO*

Аб'ект даследавання: лічбавыя фатаграфіі раслін *Arabidopsis thaliana* L. Heynh., культура раслін *Arabidopsis thaliana* L. Heynh.

Мэта даследавання: распрацоўка праграмнага забеспячэння, якое дазваляе аўтаматызаваць прадобработку і аналіз біялагічных фенатыпічнае дадзеных для расце культуры *Arabidopsis thaliana* L. Heynh.

Метады даследавання: мова праграмавання Python, метады машыннага навучання і пабудовы прадказательных мадэляў, тэхніка працы з культурай *Arabidopsis thaliana*, атрыманне малюнкаў раслін *Arabidopsis thaliana* L. Heynh. у культуры *in vitro*, морфометрычныя метады вызначэння роставых параметраў, аналіз раслін з дапамогай распрацаванага праграмнага забеспячэння.

Атрыманыя вынікі: распрацаваны праграмны комплекс для феномену аналізу раслін *Arabidopsis thaliana* L. Heynh. на аснове бібліятэк камп'ютэрнага гледжання і машыннага навучання. Распрацаваны комплекс дазваляе вымяраць плошчы паверхняў органаў раслін, а таксама прадказваць іх сырую і сухую масу. З выкарыстаннем комплексу прааналізавана выбарка з 35 кубкаў з культурай *in vitro* маладых раслін *Arabidopsis thaliana*, вырашчаных у кантрольных умовах і ва ўмовах акісляльнага стрэсу. Вымераныя значэння плошчаў каранёў і лісця па дадзеных, вынятых на аснове малюнкаў, што атрымліваюцца штодня на працягу ўсяго перыяду культывавання. Гэтыя параметры ў далейшым выкарыстаны для прадказання мас каранёў і лісця з удзелаў навучанай мадэлі, дакладнасць якой склала 96,6% па метадазе R-квадрата. У выніку зроблены вылічэнні долі мас каранёў і лісця ў адносінах да агульнай масе расліны, суадносіны плошчы каранёў і лісця да іх сухім масам, а таксама змяненне гэтых параметраў на працягу перыяду культывавання. На аснове дадзеных фенаменальнага аналізу паказана, што акісляльных стрэс, які выклікаецца сумесямі, генеральнымі гідраксільнымі радыкалы, выклікае статыстычна-пэўнае зніжэнне хуткасці набору біямасы каранёў і лісця раслін *Arabidopsis thaliana* L. Heynh.