

Определение пигментного состава зерен ячменя при помощи сверточных нейронных сетей

Артеменко Н.В.¹, Бусов И.Д.¹, Заварзин Е.А.^{1*}, Ивлева А.И.³, Кожекин М.В.^{1,2},
Глаголева А.Ю.², Генаев К.А.⁴, Генаев М.А.^{1,2}, Коваль В.С.², Комышев Е.Г.^{1,2}.

¹ Новосибирский Государственный Университет, Новосибирск, Россия;

² Институт Цитологии и Генетики СО РАН, Новосибирск, Россия

³ Самарский федеральный исследовательский центр РАН, Институт проблем управления сложными системами РАН, Самара, Россия

⁴ Exrasoft, Новосибирск, Россия

*zavarzinevg@gmail.com



SBB-2021

13th International young scientists school

Systems Biology and Bioinformatics

4-9 October 2021, Novosibirsk, Russia



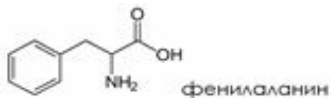
PlantGen School 2021

Young scientists school

Plant genetics, genomics, bioinformatics and biotechnology

4-9 October 2021, Novosibirsk, Russia

Исследуемые пигменты, их свойства и функции



Антоцианы

Меланины



Антоцианы:

- участвуют в регуляции роста и развитии растений;
- повышают устойчивость растений к биотическому и абиотическому стрессам;
- являются компонентами функционального питания человека (противораковый, адаптогенный, противовирусный и др. эффекты).

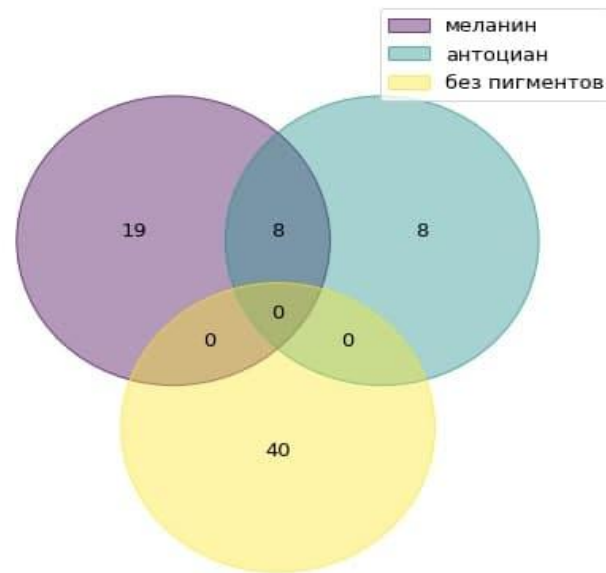
Меланин:

- связан с адаптацией растений к неблагоприятным условиям среды;
- обладает механической прочностью, химической инертностью, полупроводниковыми свойствами;
- является компонентом функционального питания человека (противораковый, антистрессовый, метаболический и др. эффекты).

Можно ли при помощи компьютерного зрения достаточно точно определить пигментацию зерен?

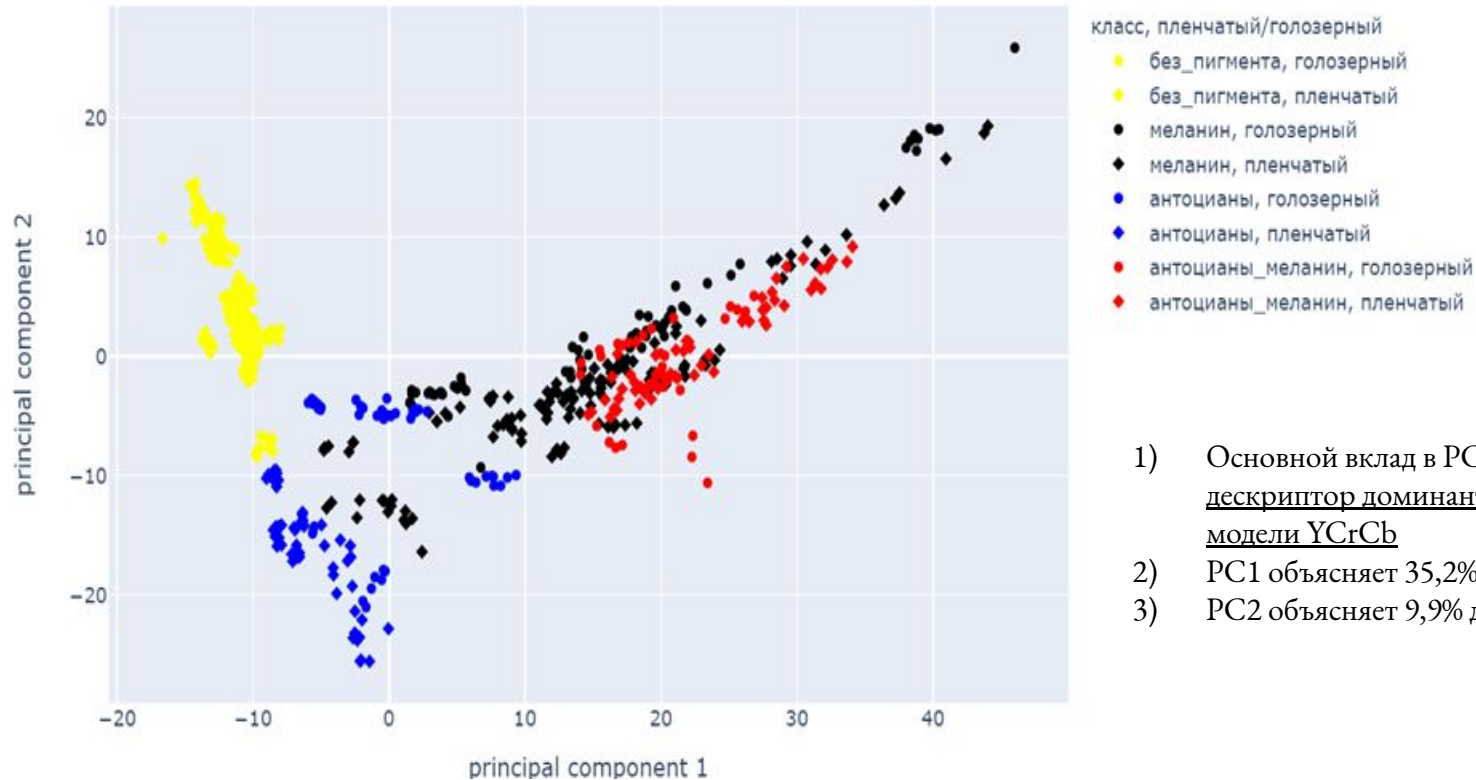
Данные проекта

- Данные фенотипирования химическими методами и методами генотипирования для 79 + 29 образцов ячменя
- 790 + 290 изображений образцов ячменя (10 изображений для каждого образца), разрешение 18 Мп
- Ручная разметка пигментации зерен и границ чаши Петри для 212 изображений (59 из 79 образцов)



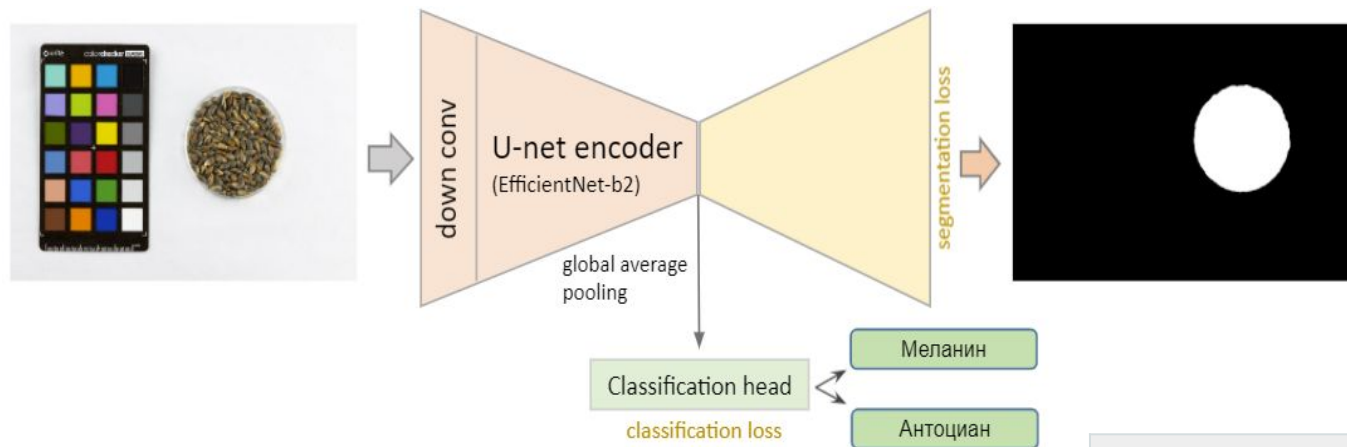
Цель проекта - создать модель компьютерного зрения для предсказания наличия пигментов (меланина и/или антоцианов) в оболочках зерна.

Метод главных компонент



- 1) Основной вклад в PC1 - средний цвет и дескриптор доминантного цвета в цветовой модели YCrCb
- 2) PC1 объясняет 35,2% дисперсии
- 3) PC2 объясняет 9,9% дисперсии

Сравнение точности моделей



- Энкодер - архитектура EfficientNet-b2
- Loss-функция - комбинация
 - Бинарной кросс-энтропии BCEWithLogitsLoss с уменьшением по среднему значению и
 - DiceLoss с функцией активации Sigmoid
- Оптимизатор Adam с начальной скоростью обучения $lr=1e-4$
- Длительность обучения 10 эпох, размер батча 2
- Сеть обучалась на RGB изображениях, трансформированных до размера 640x640 px с сохранением пропорций.

Модель	Accuracy
Random forest на отобранных дескрипторах	0.652
Классификация пигментного состава на основе CNN	0.817
Сегментации с головой для классификации	<u>0.852</u>
2-канальная сегментация	0.845

Спасибо за внимание!

Благодарности

- Математическому центру в Академгородке за организацию проекта
- Всероссийскому институту генетических ресурсов растений имени Н.И. Вавилова за предоставленный биологический материал

Буду рад ответить на вопросы и замечания!

Контакты докладчика (Заварзин Евгений):

zavarzinevg@gmail.com

<https://t.me/Evgenius2020>

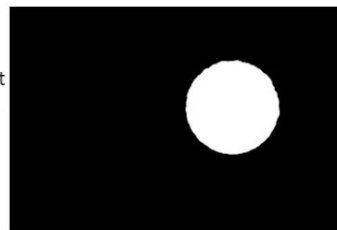


<https://habr.com/ru/post/575842/>

Оригинальное изображение



Бинаризованное изображение



U-net

Маскированное изображение



Цветовые дескрипторы

