

# Система компьютерного зрения для извлечения количественных характеристик проростков пшеницы

Бусов И. Д.  
студент 4 курса (группы 18121) бакалавриата  
Механико-математического факультета

Научный руководитель: Генаев М.А. к.б.н.

Новосибирский государственный университет

20 мая 2022

Классические методы анализа количественных характеристик проростков пшеницы на основе визуальной или тактильной экспертной оценке часто обладают рядом недостатков: субъективность, трудозатратность. В работе был реализован альтернативный подход, который основан на анализе изображений методами компьютерного зрения.



**Рис. 1:** Примеры предоставленных изображений проростков пшеницы.

**Цель.** Цель выпускной квалификационной работы – создание системы компьютерного зрения для извлечения количественных характеристик проростков пшеницы.

## Задачи:

- Разработать модели сегментации изображения растения.

**Цель.** Цель выпускной квалификационной работы – создание системы компьютерного зрения для извлечения количественных характеристик проростков пшеницы.

## Задачи:

- Разработать модели сегментации изображения растения.
- Разработать алгоритм извлечения дескрипторов (признаков) изображений растений, и с помощью классических статистических методов исследовать распределение этих дескрипторов.

**Цель.** Цель выпускной квалификационной работы – создание системы компьютерного зрения для извлечения количественных характеристик проростков пшеницы.

## Задачи:

- Разработать модели сегментации изображения растения.
- Разработать алгоритм извлечения дескрипторов (признаков) изображений растений, и с помощью классических статистических методов исследовать распределение этих дескрипторов.
- Разработать компьютерную систему для классификации проростков пшеницы по количественным характеристикам на основе методов машинного обучения.

# Протокол съемки и тестирование системы компьютерного зрения

Разработанная система компьютерного зрения была протестирована на двух задачах: определение плоидности растения и определения генотипа L-25/1102.

Для определения плоидности после фильтрации данных было 84 образца снятых с 4 различных ракурсов дважды. Промежуток между первой и второй съемкой составил 16 дней плюс минус 2 дня. Изображения снятые в первый день будем называть первой временной точкой, а во второй день второй временной точкой.

Для определения принадлежности к генотипу L-25/1102 было предоставлено 60 образцов. Каждый образец был снят лишь с одного ракурса. У данных было 2 временные точки.

Изображения, которые подаются на вход разработанной системе компьютерного зрения, были сняты по следующему протоколу: На темном фоне помещался горшок с проростком пшеницы, а сбоку прикреплялся один и тот же колорчекер. Колорчек – это цветовая палитра фиксированного размера, которая нужна для масштабирования и цветокоррекции.

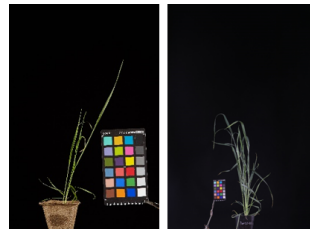


Рис. 2: Примеры изображений предоставленных для определения плоидности (слева), и для определения генотипа 1102/L-25 (справа).

Метрика для оценки точности сегментации:  
число Жаккара (Intersection over Union, IoU).

$$IoU(m_1, m_2) = \frac{|m_1 \cap m_2|}{|m_1 \cup m_2|},$$

где  $m_1 \cap m_2 = \{(i, j) | m_{2i,j} = m_{1i,j} \wedge m_{1i,j} \neq 0\}$ , а  
 $m_1 \cup m_2 = \{(i, j) | m_{1i,j} \neq 0 \vee m_{2i,j} \neq 0\}$

Метрика для оценки точности классификации :  
**accuracy.**

$$accuracy(\Omega_1, \Omega_2) = \frac{|\{\omega | \omega = \omega_1 \wedge \omega = \omega_2 \wedge \omega_1 \in \Omega_1 \wedge \omega_2 \in \Omega_2\}|}{|\Omega_1|}, \text{ где } \Omega_1, \Omega_2 \subset \mathbb{N}$$

Сегментация изображения, которое в этой работе представляется в виде упорядоченной тройки матриц, описывающих интенсивность цветовых каналов ( $R, G, B$ ), заключается в определении маски, описывающей расположение проростка. **Маской**  $m_{w,h}$  размера  $w \times h$  будем называть матрицу размера  $w \times h$ , элементы которой либо 0, либо 1.

Для решения задачи сегментации были сформированы выборки изображений следующим образом: тренировочная (33), валидационная (9) и тестовая (9), для которых маски были размечены вручную с помощью программы labelme.

Разработаны 4 модели сегментации:

- 1 На основе сверточной нейронной сети архитектуры U-Net.
- 2 На основе модифицированный индекс tgi.
- 3 На основе модифицированного индекса tgi с одним контуром.
- 4 Комбинируя 3 предыдущие.

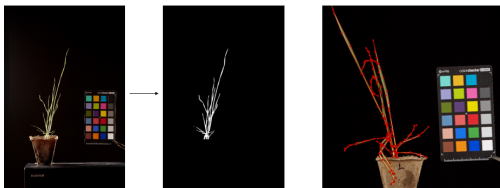


Рис. 3: Пример найденной с помощью лучшей модели сегментации маски изображения. И пример разметки изображения с помощью программы labelme



$$\theta((R, G, B)) = rR + G + bB$$

$$\tau(A)_{i,j} = \begin{cases} 1, & a_{i,j} \geq p \\ 0, & \text{иначе} \end{cases}$$

$$\max_{b,r,p} \text{IoU}(\tau(\theta(IM)), \phi(IM))$$

С одним контуром:

$$\theta((R, G, B)) = rR + G + bB$$

$$\tau(A)_{i,j} = \begin{cases} 1, & a_{i,j} \geq p \\ 0, & \text{иначе} \end{cases}$$

$$\chi(M)_{i,j} = \begin{cases} 1, & \text{если } m_{i,j} \text{ лежит в контуре проростка} \\ 0, & \text{иначе} \end{cases}$$

$$\max_{b,r,p} \text{IoU}(\chi(\tau(\theta(IM))), \phi(IM))$$

где  $b, r, p \in \mathbb{R}$  - параметры метода.

где  $IM$  - одно большое изображение, полученное путем объединения всех изображений из обучающей выборки.

К.Кеннети и Р. Шторном 1997 год[4].

# Результаты сегментации для изображений, предоставленных для определения плоидности

Модель (и выборка)	тренировочная подвыборка (IoU)	валидационная подвыборка (IoU)	тестовая подвыборка (IoU)
U-Net	0.7024	0.6472	0.7060
Модифицированный индекс tgi с одним контуром	0.743	0.7498	0.8292
Модифицированный индекс tgi	0.7914	0.7317	0.8468
Модифицированный индекс tgi с одним контуром + U-Net	0.7501	0.7003	0.7513
Модифицированный индекс tgi + U-Net	0.7573	0.6381	0.7519
<b>Комбинация моделей</b>	<b>0.7907</b>	<b>0.8261</b>	<b>0.8464</b>
Комбинация моделей + Модифицированный индекс tgi один контур	0.7766	0.8217	0.8552

Таблица 1: Таблица с результатами сегментации выборки предназначенной для определения плоидности.

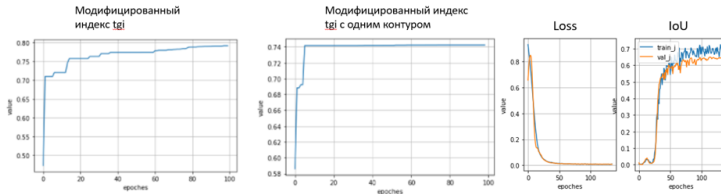
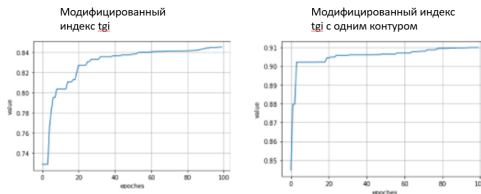


Рис. 4: График зависимости значения целевой функции лучшей особи от поколения для задачи определения плоидности. И график зависимости Loss и IoU от эпохи для U-Net. Синий на тренировочной подвыборке, оранжевый на валидационной.

# Результаты сегментации для изображений, предоставленных для определения генотипа 1102/L-25

Модель Модель (и выборка)	тренировочная подвыборка (IoU)	тестовая подвыборка(IoU)
U-Net	0.6696	0.6625
Модифицированный индекс tgi с одним контуром	0.9024	0.9100
Модифицированный индекс tgi	0.8461	0.8439
Модифицированный индекс tgi с одним контуром + U-Net	0.8903	0.8817
Модифицированный индекс tgi + U-Net	0.8265	0.8205
Комбинация моделей	0.9120	0.8946
<b>Комбинация моделей +Модифицированный индекс tgi один контур</b>	<b>0.9172</b>	<b>0.8999</b>

**Таблица 2:** Таблица с результатами сегментации выборки предназначенной для определения генотипа L-25/1102.



**Рис. 5:** График зависимости значения целевой функции лучшей особи от поколения для задачи определения генотипа 1102/L-25.

Пусть дано отсегментированное изображение  $im_{w,h} \in I$ . Для того чтобы количественно охарактеризовать форму и размер растения необходимо построить отображение  $im_{w,h}$  в элементы пространства  $\mathbb{R}^m$ .

Таким образом, одной из подзадач поставленных в этой работе, являлась нахождение некоторого семейства отображений  $\{\lambda_1, \dots, \lambda_m\}$  таких что  $\lambda_i : I \rightarrow R$ .

$\lambda_1(im_{w,h}), \dots, \lambda_m(im_{w,h})$  будем называть дескрипторами изображения  $im_{w,h}$ , которые в дальнейшем могут быть использованы для классификации методами машинного обучения.

# Набор дескрипторов для описания формы и размера побега

$\lambda_1(im_{w,h})$  – биомасса (количество элементов равных 1 в маске изображения)  $im_{w,h}$ .  
 $\lambda_2(im_{w,h})$  и  $\lambda_3(im_{w,h})$  – длина и ширина прямоугольника, описывающего изображение проростка пшеницы на фотографии.

$\lambda_4(im_{w,h})$  – линия кустистости. Линия кустистости — это медиана количества единиц в строке скелета маски изображения, которое вписано в прямоугольник описывающий проросток.

$\lambda_5(im_{w,h})$  – линия значений кустистости. Линия значений кустистости — это медиана количества единиц в строке скелета маски изображения, которое вписано в прямоугольник описывающий проросток, и если несколько таких строк имели одинаковое количество элементов не равных 0, то при подсчете медианы учитывалась лишь одна строка с уникальным количеством единиц.

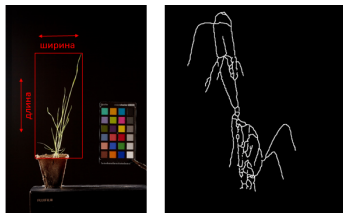


Рис. 6: Визуализация длины и ширины описанного прямоугольника. И пример скелета маски изображения, полученного с помощью библиотеки skeletonize[8].

# Набор дескрипторов для описания формы и размера побега

$\lambda_6(im_{w,h})$  – длина проростка, полученная из графа изображения.

$\lambda_7(im_{w,h})$  – количество листьев и корней в графе проростка изображения.

$\lambda_8(im_{w,h})$  – количество перегибов растения.

Таким образом, было извлечено 8 дескрипторов, позволяющих количественно описать форму и размер проростка пшеницы на изображении.



**Рис. 7:** Визуализация дескрипторов, полученных из графа изображения. Красная линия – это путь, определяющий длину. Круги – это точки перегиба и листья с корнями.

Для каждого класса изображений проростков было получено распределение по каждому дескриптору. Оценка нормальности распределений проводилась на основе  $K^2$ -теста Д'Агостина[7]. Затем строились 2 критерия. Если гипотеза о нормальности отвергалась, то для проверки равенства распределений классов использовали непараметрические критерии Манна-Уитни[5] и Краскала-Уоллиса[6]. В противном случае использовали критерий Бартлетта[8] и критерий Стьюдента[9].



Рис. 8: Блок схема для статистического анализа

# Результаты статистического анализа для задачи определения плоидности.

Дескриптор (и подвыборка)	проверка на нормальность для L-25 (p-value)	проверка на нормальность для 1102 (p-value)	критерий 1 (p-value)	критерий 2 (p-value)
Количество перегибов (1)	0.1264	0.5362	<b>0.0659</b>	0.9094
Количество листьев и корней (1)	0.4739	0.2023	<b>0.0445</b>	0.1530
Длина описывающего прямоугольника (разница)	0.0865	0.5601	0.2177	<b>0.0797</b>

**Таблица 3:** Таблица с результатами статистического анализа для выборки предназначенной для определения плоидности по тем дескрипторам и в тех выборках, где были найдены статистически значимые различия, или близкие к статистически значимым.

Дескриптор (и подвыборка)	проверка на нормальность для L-25 (p-value)	проверка на нормальность для 1102 (p-value)	критерий 1 (p-value)	критерий 2 (p-value)
Длина (разница)	0.8975	0.9038	1.0356e-10	<b>0.1262</b>
Линия кустистости (разница)	0.6060	0.5761	0.0004	<b>0.45916</b>
Линия значений кустистости (разница)	0.7215	0.4129	0.0040	<b>0.1400</b>
Ширина описывающего прямоугольника (разница)	0.3761	0.1485	0.0002	<b>0.8071</b>
Длина описывающего прямоугольника (1)	0.4128	0.8200	0.0033	<b>0.2713</b>

**Таблица 4:** Таблица с результатами статистического анализа для выборки предназначенной для определения гено типа L-25/1102 по тем дескрипторам и в тех выборках, где не были найдено статистически значимых различий.



Для получения моделей классификации были использованы методы:

- Сверточная нейронная сеть архитектуры ResNet с 18, 50 и 101 слоями.



Рис. 9: Пример изображений разных классов

Для получения моделей классификации были использованы методы:

- Сверточная нейронная сеть архитектуры ResNet с 18, 50 и 101 слоями.
- Логистическую регрессию.



Рис. 9: Пример изображений разных классов

Для получения моделей классификации были использованы методы:

- Сверточная нейронная сеть архитектуры ResNet с 18, 50 и 101 слоями.
- Логистическую регрессию.
- Алгоритм машинного обучения случайный лес.



Рис. 9: Пример изображений разных классов

Для получения моделей классификации были использованы методы:

- Сверточная нейронная сеть архитектуры ResNet с 18, 50 и 101 слоями.
- Логистическую регрессию.
- Алгоритм машинного обучения случайный лес.
- Комбинируя 3 предыдущие.



Рис. 9: Пример изображений разных классов

# Результаты классификации по количественным характеристикам

Модель (и выборка)	тренировочная подвыборка (ассигасу)	тестовая подвыборка(ассигасу)
ResNet(все)	0.4655	0.5385
Логистическая регрессия(все)	0.7069	0.6154
<b>Случайный лес(все)</b>	<b>0.7586</b>	<b>0.6538</b>
Комбинация моделей(все)	0.7414	0.5770
ResNet(1)	0.5517	0.5385
Логистическая регрессия(1)	0.6380	0.5770
<b>Случайный лес(1)</b>	<b>0.6552</b>	<b>0.6154</b>
Комбинация моделей(1)	0.6380	0.6539
ResNet(2)	0.5345	0.46154
Логистическая регрессия(2)	0.5345	0.4615
<b>Случайный лес(2)</b>	<b>0.8103</b>	<b>0.7692</b>
Комбинация моделей(2)	0.5345	0.4615
Логистическая регрессия(разница)	0.6552	0.4615
<b>Случайный лес(разница)</b>	<b>0.6900</b>	<b>0.4615</b>

Таблица 5: Таблица с результатами классификации по плоидности.

Модель (и выборка)	тренировочная подвыборка (ассигасу)	тестовая подвыборка(ассигасу)
ResNet(все)	0.95	0.95
Логистическая регрессия(все)	0.98	1.0
<b>Случайный лес(все)</b>	<b>1.0</b>	<b>1.0</b>
Комбинация моделей(все)	0.99	1.0
ResNet(1)	0.96	1.0
Логистическая регрессия(1)	0.98	1.0
<b>Случайный лес(1)</b>	<b>1.0</b>	<b>1.0</b>
Комбинация моделей(1)	1.0	1.0
ResNet(1)	1.0	1.0
Логистическая регрессия(2)	0.98	1.0
<b>Случайный лес (2)</b>	<b>1.0</b>	<b>1.0</b>
Комбинация моделей(2)	1.0	1.0
Логистическая регрессия(разница)	0.88	1.0
<b>Случайный лес(разница)</b>	<b>1.0</b>	<b>1.0</b>

Таблица 6: Таблица с результатами классификации по принадлежности к генотипу 1102/L-25

- Разработаны модели сегментации изображения проростка растения на основе индекса  $tg i$ , нейронных сетей и комбинации этих методов. Комбинация методов дает наилучший результат (для задачи определения плоидности – 0.8464 (IoU), для задачи определения генотипа 1102/L-25 – 0.8999 (IoU) на тестовых выборках).

- Разработаны модели сегментации изображения проростка растения на основе индекса  $tg\alpha$ , нейронных сетей и комбинации этих методов. Комбинация методов дает наилучший результат (для задачи определения плоидности – 0.8464 (IoU), для задачи определения генотипа 1102/L-25 – 0.8999 (IoU) на тестовых выборках).
- Разработан алгоритм для извлечения 8 дескрипторов, описывающих форму и размер проростка на изображении.

- Разработаны модели сегментации изображения проростка растения на основе индекса  $tgi$ , нейронных сетей и комбинации этих методов. Комбинация методов дает наилучший результат (для задачи определения плоидности – 0.8464 (IoU), для задачи определения генотипа 1102/L-25 – 0.8999 (IoU) на тестовых выборках).
- Разработан алгоритм для извлечения 8 дескрипторов, описывающих форму и размер проростка на изображении.
- Проведен статистический анализ распределения дескрипторов изображений проростков. Статистически значимых различий в распределении дескрипторов изображений проростков разной плоидности, за редким исключением, не обнаружено, но для изображений проростков разных генотипов 1102/L-25 почти везде были найдены статистически значимые различия в распределениях извлеченных дескрипторов.



- Разработаны модели сегментации изображения проростка растения на основе индекса  $tgi$ , нейронных сетей и комбинации этих методов. Комбинация методов дает наилучший результат (для задачи определения плоидности – 0.8464 (IoU), для задачи определения генотипа 1102/L-25 – 0.8999 (IoU) на тестовых выборках).
- Разработан алгоритм для извлечения 8 дескрипторов, описывающих форму и размер проростка на изображении.
- Проведен статистический анализ распределения дескрипторов изображений проростков. Статистически значимых различий в распределении дескрипторов изображений проростков разной плоидности, за редким исключением, не обнаружено, но для изображений проростков разных генотипов 1102/L-25 почти везде были найдены статистически значимые различия в распределениях извлеченных дескрипторов.
- Разработана компьютерная система для классификации проростков пшеницы по целевому признаку на основе методов машинного обучения (логистическая регрессия, случайный лес, нейронные сети архитектуры ResNet) и их комбинации. Результаты лучших моделей классификации для решения задачи определения плоидности на тестовой выборке были следующими: без разделения на временные точки – 0.654 (accuracy), первая временная точка – 0.615 (accuracy), вторая временная точка – 0.769 (accuracy), разница – 0.462 (accuracy). Лучшая модель классификации по генотипу 1102/L-25 предсказывала генотип безошибочно.



Ronneberger O. et al., U-net: Convolutional networks for biomedical image segmentation, 2015 //arXiv preprint arXiv:1505.04597.



He, Kaiming; Zhang, Xiangyu; Ren, Shaoqing - Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE. pp. 770–778 //arXiv:1512.03385



Alt V.V., Pestunov I.A., Melnikov P.V., Elkin O.V Avtomatizirovannoe obnaruzhenie sornyakov i otsenka kachestva vskhodov sel'skokhozyaistvennykh kul'tur po RGB-izobrazheniyam [Automated detection of weeds and evaluation of crop sprouts quality based on RGB images]. Sibirskii vestnik sel'skokhozyaistvennoi nauki [Siberian Herald of Agricultural Science], 2018, vol. 48, no. 5 , pp. 52–60. DOI: 10.26898/0370-8799-2018-5-7



By Kenneth Price and Rainer Storn, April 01, 1997.<https://www.drdobbs.com/database/differential-evolution/184410166>



H.B. Mann and D.R. Whitney , “On a test of whether one of two random variables is stochastically larger than the other”, The Annals of Mathematical Statistics, 1947, Vol. 18, pp. 50-60.



W. H. Kruskal W. W. Wallis, “Use of Ranks in One-Criterion Variance Analysis”, Journal of the American Statistical Association, 1952, Vol. 47, Issue 260, pp. 583-621.



D'Agostino, R. B., “An omnibus test of normality for moderate and large sample size”, Biometrika, 1971, 58, 341-348.



Bartlett, M. S. . Properties of Sufficiency and Statistical Tests. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences,1937, Vol. 160, No.901, pp. 268-282.



Yuen, Karen K., and W. J. Dixon. “The Approximate Behaviour and Performance of the Two-Sample Trimmed t.” Biometrika, Accessed 30 Mar. 2021, vol. 60, no. 2, 1973, pp. 369-374. JSTOR, [www.jstor.org/stable/2334550](http://www.jstor.org/stable/2334550).



T. Hastie, R. Tibshirani and J. Friedman. "Elements of Statistical Learning"Springer, 2009.



Hsiang-Fu Yu, Fang-Lan Huang, Chih-Jen Lin (2011). Dual coordinate descent methods for logistic regression and maximum entropy models. Machine Learning 85(1-2):41-75.



Haykin S. .Neural Networks and Learning Machines. Pearson Education. Inc., Upper Saddle River, New Jersey, 2009.



Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. introduction to Algorithms. MIT Press. MIT Electrical Engineering and Computer Science. MIT Press. 18 June 1990. ISBN 978-0-262-03141-7. Retrieved 2017-07-02.



Olaf Ronneberger, Philipp Fischer, and Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation,, 18 May 2015 <https://arxiv.org/pdf/1505.04597.pdf>.



Kentaro Wada, Michael Pitidis, Hussein Abdulwahid, 2018, <https://github.com/wkentaro/labelme>.



Буэно, Суарес, Эспиноса. Обработка изображений с помощью OpenCV = Learning Image Processing with OpenCV. — М.: ДМК-Пресс, 2016. — 210 с. — ISBN 978-5-97060-387-1.



T.-C. Lee, R.L. Kashyap and C.-N. Chu, Building skeleton models via 3-D medial surface/axis thinning algorithms. Computer Vision, 1994 Graphics, and Image Processing, 56(6):462-478.



Mingxing Tan, Quoc V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 28 May 2019, <https://arxiv.org/abs/1905.11946>

# Система компьютерного зрения для извлечения количественных характеристик проростков пшеницы

Бусов И. Д.  
студент 4 курса (группы 18121) бакалавриата  
Механико-математического факультета

Научный руководитель: Генаев М.А. к.б.н.

Новосибирский государственный университет

20 мая 2022

**Маской**  $m_{w,h}$  размера  $w \times h$  будем называть матрицу размера  $w \times h$ , элементы которой либо 0, либо 1.

Введем **оператор наложения матриц**  $\times_1 : A \times A \longrightarrow A$  такой что  $a, b, c \in A$   
 $a \times_1 b = c \iff a_{i,j} * b_{i,j} = c_{i,j}$ , где  $A$  - множество всех матриц фиксированного размера.

И **оператор масочного объединения матриц**  $\times_2 : A \times A \longrightarrow A$  такой что  $a, b, c \in A$   
 $a \times_2 b = c \iff \min(a_{i,j} + b_{i,j}, 1) = c_{i,j}$ , где  $A$  - множество всех матриц фиксированного размера.

Под **метрикой для задачи сегментацией** в данной работе будем подразумевать некоторый функционал  $\mu : M \times M \longrightarrow \mathbb{R}$ , где  $M$  - множество масок.

Под **метрикой для задачи классификации** в данной работе будет подразумеваться некоторый функционал  $\zeta : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$ , где  $\Omega_1, \Omega_2 \subset \mathbb{N}$ .

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \times_1 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \times_2 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Рис. 10: Введенные операции на примере.

**Изображением**  $im_{w,h}$  размера  $w \times h$  мы будем называть упорядоченную тройку  $(R, G, B)$ , где  $R = (r_{i,j})$ ,  $G = (g_{i,j})$ ,  $B = (b_{i,j})$  - матрицы размера  $w \times h$ ,  $w$  - ширина изображения, а  $h$  - длина изображения. А также элементы матриц должны удовлетворять следующим условиям:

$$r_{i,j} \in \mathbb{N} \wedge 0 \leq r_{i,j} \leq 255$$

$$g_{i,j} \in \mathbb{N} \wedge 0 \leq g_{i,j} \leq 255$$

$$b_{i,j} \in \mathbb{N} \wedge 0 \leq b_{i,j} \leq 255, \text{ где } i = \overline{1, h} \ j = \overline{1, w}$$

**Маской**  $m_{w,h}$  размера  $w \times h$  будем называть матрицу размера  $w \times h$ , элементы которой либо 0, либо 1.

Пусть существует отображение из множества изображений  $I$  в множество масок  $M$   $\phi : I \longrightarrow M$ . **Задачей сегментацией** будем называть поиск отображения  $\phi$ .

Пусть существует отображение из множества изображений  $I$  в некоторое подмножество натуральных чисел  $\Omega \subset \mathbb{N}$  и пусть существует отображение  $\psi : I \longrightarrow \Omega$ . **Задачей классификации** будем называть поиск отображения  $\psi$ . Если  $|\Omega| = 2$ , то будем называть такую задачу классификации - задачей бинарной классификации.

Под **метрикой для задачи сегментацией** в данной работе будем подразумевать некоторый функционал  $\mu : M \times M \longrightarrow \mathbb{R}$ , где  $M$  - множество масок.

Под **метрикой для задачи классификации** в данной работе будет подразумеваться некоторый функционал  $\zeta : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{R}$ , где  $\Omega_1, \Omega_2 \subset \mathbb{N}$ .

Будем говорить, что мы решили задачу сегментации поиска функции  $\phi : I \longrightarrow M$  с точность асс по метрики  $\mu$ , если мы нашли такое отображение  $f : I \longrightarrow M$ , что  $асс = \mu(\phi(im_{w,h}), f(im_{w,h}))$ ,  $I$ -множество изображений,  $im_{w,h} \in I$ , а  $M$  - множество масок. Найденное отображение  $f$  будем называть моделью сегментации, а  $\phi(im_{w,h})$  будем называть маской изображения  $im_{w,h}$ .

Будем говорить, что мы решили задачу классификации поиска функции  $\psi : I \longrightarrow \Omega$  с точность асс по метрики  $\zeta$ , если мы нашли такое отображение  $f : I \longrightarrow \Omega$ , что  $асс = \zeta(\psi(I), f(I))$ ,  $I$ -множество изображений, а  $\Omega$  - подмножество натуральных чисел, такое что  $Im(\psi) \subseteq \Omega$ . Найденное отображение  $f$  будем называть моделью классификации.

Введем **оператор наложения матриц**  $\times_1 : A \times A \longrightarrow A$  такой что  $a, b, c \in A$   
 $a \times_1 b = c \iff a_{i,j} * b_{i,j} = c_{i,j}$ , где  $A$  - множество всех матриц фиксированного размера.

И **оператор масочного объединения матриц**  $\times_2 : A \times A \longrightarrow A$  такой что  $a, b, c \in A$   
 $a \times_2 b = c \iff \min(a_{i,j} + b_{i,j}, 1) = c_{i,j}$ , где  $A$  - множество всех матриц фиксированного размера.

А также введем **оператор сегментации**  $seg : \Phi \times I \longrightarrow I$ .  
 $seg(\phi, im) = (R \times_1 \phi(im), G \times_1 \phi(im), B \times_1 \phi(im))$ , где  $im \in I$ ,  $im = (R, G, B)$   $\Phi$  - множество всех отображений  $\phi : I \longrightarrow M$ .  $seg(\phi, im)$  будем называть отсегментированным при помощи модели  $\phi$  изображением  $im$ .

Будем обозначать через  $ONE_{w,h}$  - матрицу размера  $w \times h$ , все элементы которой равны 1.

**Обучающей выборкой** будем называть упорядоченную пару  
 $(X, y) = \{(X_i, y_i) | X_i \in S_1, y_i \in S_2, i = \{1, \dots, n\}\}$ , где  $S_1, S_2$  - некоторые множества.



Задача создания системы компьютерного зрения разбивается на следующие подзадачи:

- 1 Разбить выборку  $(\mathbb{X}, y)$  на  $\text{train}, \text{val}$  и  $\text{test}$ .
- 2 Решить задачу сегментации с максимальной точностью по выбранной метрике по поиску  $\phi$  отображающей изображение проростка пшеницы в маску, которая содержит информацию о расположении проростка на фотографии. Более формально, если  $(i, j)$  - упорядоченная пара индексов изображения  $im = ((r_{i,j}), (g_{i,j}), (b_{i,j}))$  является элементом множества  $\text{Shoot} = \{(i, j) | i \text{ и } j \text{ соответствуют расположению проростка на фотографии}\}$ , то  $m_{i,j} = 1$ , где  $m = \phi(im)$ , в противном случае  $m_{i,j} = 0$ . В дальнейшем  $\phi$  всегда определена одинакова, и будет опускаться.
- 3 Решить задачу классификации с максимальной точностью по выбранной метрике по поиску функции, которая отображает отсегментированное изображение с проростком пшеницы в количественную характеристику, которую можно отождествить с некоторым натуральным числом.

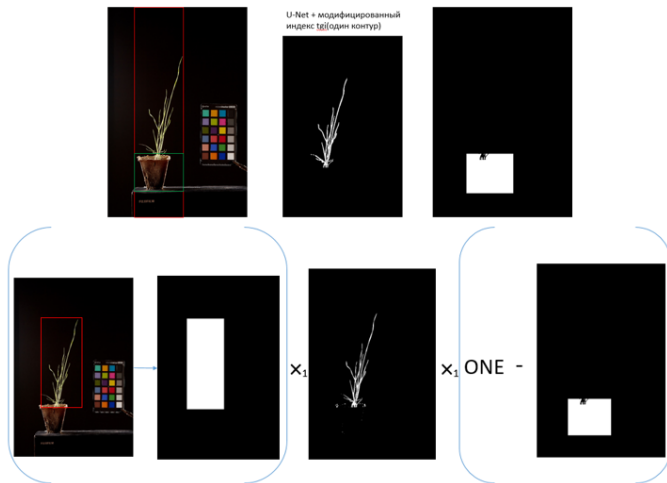


Рис. 11: Иллюстрация комбинирования моделей

ONE – это матрица все элементы которой равны 1.

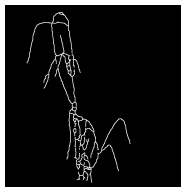
Пусть  $im_{w,h} \in I$  тогда масштабирование этого изображения было реализовано по следующему алгоритму, вначале определим  $res = (R', G', B')$ , где  $R', G', B'$  – подматрицы  $R, G, B$ , полученные вычеркиванием строк и столбцов, каждый элемент которых лежит вне прямоугольника, описывающего маску изображения  $im_{w,h}$ . Мы брали большое изображение фиксированного размера состоящие из нулей, и если  $m_{i,j}$  не был равен нулю, где  $m$  – маска изображения  $res$ , то элемент стоящей на позиции  $(i,j)$  в каждом канале изображения мы приравнивали элементу стоящему на позиции  $(i,j)$  соответствующего канала изображения  $im_{w,h}$ . Новое большое изображение и было отмасштабированным изображением  $im_{w,h}$ . Масштабирование необходимо, так как одна из простых аугментаций ResNet-ов – это сжатие изображения до размера  $512 \times 512$ .



Рис. 12: Пример отмасштабированного изображения, на котором обучались нейронные сети архитектуры ResNet

## Представление изображения в виде графа

Вначале при помощи библиотеки OpenCV[7] к маске изображения было применено небольшое размытие, это было сделано для того, чтобы уменьшить количество областей связности. После, используя библиотеку skeletonize [8], мы построили скелет изображения, то есть уменьшили количество элементов равных 1 в маске изображения таким образом, что ширина маски проростка везде равнялась единице. Скелет маски изображения представлялся в виде графа следующим образом, каждый элемент скелета который равен 1 был вершиной нашего графа. Если две вершины имели соответствующие координат  $(i_1, j_1)$ ,  $(i_2, j_2)$ , то мы говорили, что они находились рядом, если верно следующее высказывание  $|i_1 - i_2| \leq 1 \wedge |j_1 - j_2| \leq 1 \wedge (|i_1 - i_2| + |j_1 - j_2| \neq 0)$ . Любые две находящиеся рядом вершины мы соединяли ребром, если  $|i_1 - i_2| = 0 \vee |j_1 - j_2| = 0$ , то вес ребра 1, иначе вес ребра  $\sqrt{2}$ . Таким образом мы отображали изображение в взвешенный неориентированный граф  $(V_{im}, E_{im})$ , где  $V_{im}$  - множество вершин, а  $E_{im}$  - множество ребер в графе изображения  $im$ . Корнями графа изображения будем называть вершины, которые находятся ниже других вершин в своей области связности.



- **Удаление циклов.** Для извлечения длины мы работали только с наибольшей по количеству вершин областью связности. Если бы нам были предоставлены 3-х мерные изображения то циклов, за очень редким исключением, в графе растения не было бы, то есть циклы не свойственны природе проростка пшеницы, и потому в этой работе мы от них избавлялись. При поиски длины мы не удаляли циклы длины 3, так как они не противоречат биологии проростка. С помощью обхода в глубину, идя из одной вершину в отличную от нее с наименьшем приоритетом другую (если из вершины  $v_1$  можно пойти в несколько вершин  $v_i \neq v_1$ , то мы идем в такую вершину  $v_2$ , что  $\forall i \ pr(v_1, v_2) \leq pr(v_1, v_i)$ ), мы находили циклы графа. И в каждом цикле удаляли ребро между первой вершиной, которая находится ниже предшествующей в данном обходе, и предшествующей ей.

- **Удаление циклов.** Для извлечения длины мы работали только с наибольшей по количеству вершин областью связности. Если бы нам были предоставлены 3-х мерные изображения то циклов, за очень редким исключением, в графе растения не было бы, то есть циклы не свойственны природе проростка пшеницы, и потому в этой работе мы от них избавлялись. При поиски длины мы не удаляли циклы длины 3, так как они не противоречат биологии проростка. С помощью обхода в глубину, идя из одной вершину в отличную от нее с наименьшем приоритетом другую (если из вершины  $v_1$  можно пойти в несколько вершин  $v_i \neq v_1$ , то мы идем в такую вершину  $v_2$ , что  $\forall i \text{ } pr(v_1, v_2) \leq pr(v_1, v_i)$ ), мы находили циклы графа. И в каждом цикле удаляли ребро между первой вершиной, которая находится ниже предшествующей в данном обходе, и предшествующей ей.
- **Подсчет длины.** Для более корректного определения корней графа изображения при подсчете длины удалялись все вершины, соответствующие элементам маски изображения, которые находились ниже наивысшего элемента маски горшка не равного 0. Далее с помощью алгоритма Дейкстры для каждой корневой вершины находили расстояния от нее до всех остальных вершин. И определяли длину растения как наименьшее расстояние среди наибольших расстояний от корневых вершин до всех остальных вершин графа.

- **Удаление циклов.** Циклы удалялись по тому же алгоритму, что и при подсчете длины, только при поиски количества листьев и корней удалялись все циклы, в том числе и циклы длины 3.

- **Удаление циклов.** Циклы удалялись по тому же алгоритму, что и при подсчете длины, только при поиски количества листьев и корней удалялись все циклы, в том числе и циклы длины 3.
- **Нахождение листьев в графе.** С помощью обхода в глубину с использованием функции приоритета вершин в ацикличном графе были найдены корни (при поиски количества листьев и корней, вершины, соответствующие элементам маски изображения, которые находились ниже наивысшего элемента маски горшка не равного 0, не удалялись из графа) и листья удовлетворяющие следующему условию: 21 (число найденное эмпирическим путем) предков листа должны иметь не более 3 смежных вершин. Это условие нужно для того, чтобы избавиться от ошибок сегментации и проблем связанных с 2-х мерностью изображения.



- **Удаление циклов.** Циклы удалялись по тому же алгоритму, что и при подсчете длины, только при поиски количества листьев и корней удалялись все циклы, в том числе и циклы длины 3.
- **Нахождение листьев в графе.** С помощью обхода в глубину с использованием функции приоритета вершин в ациклическом графе были найдены корни (при поиски количества листьев и корней, вершины, соответствующие элементам маски изображения, которые находились ниже наивысшего элемента маски горшка не равного 0, не удалялись из графа) и листья удовлетворяющие следующему условию: 21 (число найденное эмпирическим путем) предков листа должны иметь не более 3 смежных вершин. Это условие нужно для того, чтобы избавиться от ошибок сегментации и проблем связанных с 2-х мерностью изображения.
- **Подсчет реального количества листьев и корней графа.** При подсчете количества листьев и корней мы не учитывали того факта, что в графе могут быть несколько областей связности, с учетом областей связности реальное количество листьев и корней равняется найденному количеству листьев и корней минус удвоенное количество областей связности минус 2.

- **Создание нового графа.** Для подсчета количества перегибов в данной работе составлялся новый граф из ребер и вершин, лежащих в наикратчайших путях от произвольно выбранного корня в каждой области связности до листьев и других корней.

- **Создание нового графа.** Для подсчета количества перегибов в данной работе составлялся новый граф из ребер и вершин, лежащих в наикратчайших путях от произвольно выбранного корня в каждой области связности до листьев и других корней.
- **Поиск перегибов.** В этом новом графе с помощью обхода в глубину с использованием функции приоритета вершин были найдены вершины, в которых произошла смена направления обхода вершин из направления снизу вверх в направление сверху вниз. Затем если новое направление (сверху вниз) продержалось больше 15 вершин (количество вершин найдено эмпирическим путем), то мы объявляли вершину перегибом в графе изображения. Поиск перегибов производился во всех областях связности.

# Почему не извлекались цветовой дескрипторы?

Были рассмотрены распределения медиан цветов в пространстве HSV(Hue,Saturation,Value) для изображений снятых в разное время, то есть для изображений условия съемки которых были немного различными. Везде были найдены статистически значимые различия в распределение медиан цветов как до цветокоррекции, так и после.

Таким образом, цветокоррекция улучшает результат, но все же не способна нивелировать разные условия съемки, и потому в дальнейшем цветовой дескрипторы не извлекались из изображений.

Признак	Проверка на нормальность (1-ая выборка) p-value	Проверка на нормальность (2-ая выборка) p-value	Критерий Бартлетта p-value	Критерий Стьюдента p-value
Медиана по H до цветокоррекции	0.1538122552947248	0.39221902885020227	1.3444003814998725e-06	0.00999197407784985
Медиана по H после цветокоррекции	0.25091661840037993	0.17204732454393049	7.873644957570192e-05	0.0036547249363259116
Медиана по S до цветокоррекции	0.30143520987224787	0.5192724324865079	0.019649391639021317	3.090268397431219e-17
Медиана по S после цветокоррекции	0.5241809135496223	0.11398854785408377	0.022175513005175545	4.6066330647041495e-10
Медиана по V до цветокоррекции	0.6263584467335633	0.5588911552932452	0.23620056292242003	3.7466973524324925e-13
Медиана по V после цветокоррекции	0.8554539970428207	0.4302961599803635	0.5556647661087724	5.883266882030908e-10

Таблица 7: Таблица со значениями p-value по распределениям медиан цветов.

# Выбросы при статистическом анализе

Дескриптор (и подвыборка)	количество выбросов при исследовании генотипов L-25/1102	количество выбросов при исследовании плоидности
Длина (1)	0	1
Длина (2)	3	1
Длина (разница)	3	2
Длина (все)	3	0
Количество перегибов (1)	0	4
Количество перегибов (2)	0	1
Количество перегибов (разница)	0	4
Количество перегибов (все)	0	4
Количество листьев и корней (1)	1	1
Количество листьев и корней (2)	0	1
Количество листьев и корней (разница)	0	3
Количество листьев и корней (все)	0	3
Линия кустистости (1)	0	0
Линия кустистости (2)	0	0
Линия кустистости (разница)	0	0
Линия кустистости (все)	0	0
Линия значений кустистости (1)	0	0
Линия значений кустистости (2)	0	0
Линия значений кустистости (разница)	0	0
Линия значений кустистости (все)	0	0
Ширина описывающего прямоугольника (1)	0	0
Ширина описывающего прямоугольника (2)	4	0
Ширина описывающего прямоугольника (разница)	4	0
Ширина описывающего прямоугольника (все)	0	0
Длина описывающего прямоугольника (1)	0	1
Длина описывающего прямоугольника (2)	0	1
Длина описывающего прямоугольника (разница)	1	1
Длина описывающего прямоугольника (все)	0	1
Биомасса (1)	0	0
Биомасса (2)	3	0
Биомасса (разница)	3	0
Биомасса (все)	3	0

Таблица 8: Таблица с количеством выбросов.

# Результаты классификации по плоидности при помощи сверточной нейронной сети ResNet.

Модель (и выборка)	train(accuracy)	val(accuracy)	test(accuracy)
ResNet18(все)	0.5	0.4167	0.3846
<b>ResNet50(все)</b>	<b>0.4412</b>	<b>0.5</b>	<b>0.5385</b>
ResNet101(все)	0.4412	0.5	0.5385
ResNet18(1)	<b>0.5588</b>	<b>0.5417</b>	<b>0.5385</b>
ResNet50(1)	0.5588	0.4583	0.4615
ResNet101(1)	0.4412	0.5	0.5385
ResNet18(2)	<b>0.5588</b>	<b>0.5</b>	<b>0.4615</b>
ResNet50(2)	0.6765	0.5	0.4231
ResNet101(2)	0.6176	0.5	0.4231

Таблица 9: Таблица с результатами классификации по плоидности с помощью ResNet-ов с 18, 50 и 101 слоями.

Без разделения на временные точки

Первая временная точка

Вторая временная точка

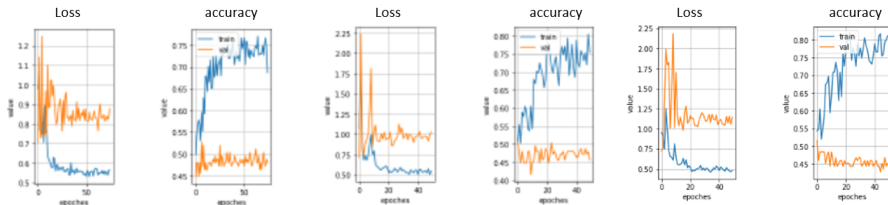


Рис. 14: График зависимости для лучших ResNet-ов, решающих задачу классификации по плоидности, Loss и accuracy от эпохи. Синий график – train, оранжевый – val

# Результаты классификации по генотипу 1102/L-25 при помощи сверточной нейронной сети ResNet

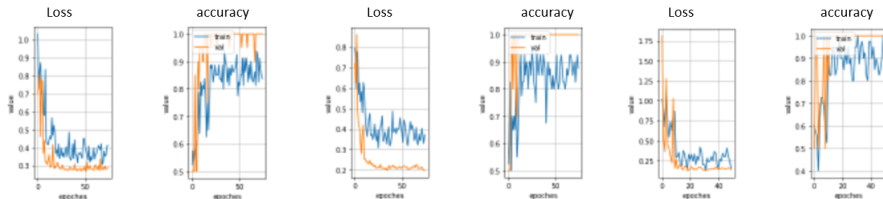
Модель (и выборка)	train(accuracy)	val(accuracy)	test(accuracy)
ResNet18(все)	0.9375	1.0000	0.9000
<b>ResNet50(все)</b>	<b>0.9375</b>	<b>1.000</b>	<b>0.950</b>
ResNet101(все)	0.8125	0.9500	0.9000
ResNet18(1)	1.000	0.950	0.9000
<b>ResNet50(1)</b>	<b>0.9500</b>	<b>1.0000</b>	<b>1.0000</b>
ResNet101(1)	1.0000	1.0000	0.8000
<b>ResNet18(2)</b>	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>
ResNet50(2)	0.9750	1.0000	1.0000
ResNet101(2)	0.9750	1.0000	1.0000

**Таблица 10:** Таблица с результатами классификации по принадлежности к генотипу L-25/1102 с помощью ResNet-ов с 18, 50 и 101 слоями.

Без разделения на временные точки

Первая временная точка

Вторая временная точка



**Рис. 15:** График зависимости для лучших ResNet-ов, решающих задачу классификации по принадлежности к генотипу 1102/L-25, Loss и accuracy от эпохи. Синий график – train, оранжевый – val