

УДК:517.534

ПОСТРОЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ ЗАДАЧ ПРОГНОЗИРОВАНИЯ БОЛЕЗНЕЙ СЕЛЬСКОХОЗЯЙСТВЕННЫХ РАСТЕНИЙ

Сабитов¹ Б.Р., Картанова² А., Алмабекова¹ Н.А., Джунушалиева¹ А.Д., Ажаакматова¹ А.А., Тенирбеди² у И., Имангазы² у. Н., Цю Ифен¹, Чжан Цайцзюй¹

¹КНУ им. Ж.Баласагына, ²КТУ им. И.Раззакова, ³КРСУ им.Б.Ельцина

В этой статье рассматривается вопрос, как управлять обучением нейронной сети в Tensorflow. Для анализа точности построенной модели использовано база данных больных и здоровых листьев кукурузы. Проведен анализ модели для различных эпох обучения. Построены ROC кривые болезней растений и получена матрица погрешностей.

Ключевые слова: нейронная сеть, модель, болезни кукурузы, распознавание, обучение, прогнозирование.

АЙЫЛ ЧАРБА ӨСҮМДҮҮЛӨРҮНҮН ООРУЛАРЫН БОЛГОН ПРОБЛЕМАЛАРЫНЫН НЕЙРАЛДЫК ТАРМАГЫН КУРУУ.

Сабитов¹ Б.Р., Картанова² А., Алмабекова¹ Н.А., Жунушалиева¹ А.Д., Ажаакматова¹ А.А., Тенирбеди² у И., Имангазы² у. Н., Цю Ифен¹, Чжан Цайцзюй¹

¹Ж. Баласагын ат. КУУ, ²И.Раззаков ат. КМТУ, ³Б.Ельцин ат. КОСУ

Бул макалада Tensorflow'та нейрондук тармакты окутууну кантип башкаруу керек деген суроо каралат. Түзүлгөн моделдин тактыгын талдоо үчүн жүгөрү жалбырактарынын оорулуу жана соолуктарынын маалымат базасы колдонулган. Модель ар кандай окутуу доорлору үчүн талданган. Өсүмдүк ооруларынын ROC ийри сызыктары түзүлүп, ката матрицасы алынды.

Баштапкы сөздөр: Нейрондук тармак, модель, жүгөрү оорулары, таануу, үйрөнүү, болжолдоо.

CONSTRUCTION OF A NEURAL NETWORK FOR PROBLEMS OF PREDICTION OF DISEASES OF AGRICULTURAL PLANTS

Sabitov¹ B.R., Kartanova² A., Almabekova¹ N.A., Dzhunushalieva¹ A.D., Azhaakmatova¹ A.A., Tenirbedi² Y., Imangazy N. Zu Ifen¹, Chjan Zayzshuy¹

¹KNU named after J. Balasagyn, ²KTU named after I. Razzakov, ³KRSU named after B. Yeltsin

This article addresses the question of how to manage neural network training in Tensorflow. To analyze the accuracy of the constructed model, a database of diseased and healthy maize leaves was used. The model was analyzed for different training epochs. ROC curves of plant diseases were constructed and an error matrix was obtained.

Key words: neural network, model, corn diseases, recognition, learning, prediction.

Введение

Болезни листьев кукурузы значительно снижают урожай кукурузы; поэтому мониторинг и выявление болезней в течение вегетационного периода имеют решающее значение. Некоторые из текущих исследований основаны на изображениях с простым фоном, а реалистичные полевые настройки полны фонового шума, что усложняет эту задачу. В данной статье изучается задача распознавания болезней кукурузы на базе данных цветных изображений красные, зеленые и синие (RGB) изображения с наших экспериментальных полей и общедоступного набора данных, и они содержат в общей сложности четыре категории, а именно: южная пятнистость листьев кукурузы (SCLB), серая пятнистость листьев (GLS), южная кукурузная ржавчина (SR) и здоровая (H). В этой статье предлагается модель, отличная от сверточных нейронных сетей (CNN), основанная увеличении эпохи при обучении нейронной сети. Результаты показывают, что наша модель достигает наилучшей производительности, а механизм внимания играет чрезвычайно важную роль. Сбор и подготовка данных. Изображения больных и здоровых листьев кукурузы загружаются из Plant Village.

Обзор научных публикаций

Кукуруза — одна из самых важных культур для человечества с самой высокой урожайностью в мире (Ranum et al., 2014, [1]). Заболевания кукурузы могут привести к серьезному снижению урожайности, что является серьезной проблемой ([2] Savary et al., 2012). Поэтому очень важно своевременно выявлять и контролировать болезни кукурузы в период вегетации. Точная идентификация болезней кукурузы затруднена для растениеводов, которые могут не быть профессионалами в области патологии растений, а экспертная идентификация требует больших затрат и времени ([3] Oupparaphan, 2017). Традиционные методы распознавания изображений и глубокое обучение постепенно входят в область распознавания болезней растений ([4] Saleem et al., 2019). В недавнем исследовании выявления болезней кукурузы сравнивались пять стандартных методов машинного обучения ([5] Panigrahi et al., 2020), а именно: наивный байесовский метод (NB), дерево решений (DT), K-ближайший сосед (KNN), SVM и случайный лес (RF), при этом RF достигает наивысшей точности 79,23%. Однако традиционное машинное обучение в основном ограничено извлечением и представлением признаков. Глубокое обучение добилось значительного прогресса в идентификации болезней растений [6] (Liu and Wang, 2021). С тех пор как AlexNet был предложен в 2012 г. ([7], Krizhevsky, A., и др (2012), сверточные нейронные сети (СНС) широко используются в задачах распознавания болезней растений. В области идентификации болезней растений [8] Dhaka et al. (2021) предоставил систематический обзор соответствующих методов глубокого обучения. Благодаря своей низкой сложности облегченная CNN достигла удовлетворительных результатов в идентификации болезней кукурузы ([3] Oupparaphan, 2017). Система на основе CNN ([9] DeChant et al., 2017) была внедрена для автоматической идентификации северной пятнистости листьев, решая проблемы ограниченности данных и

различных нарушений, появляющихся на изображениях, выращенных в полевых условиях. Ahila Priyadharshini, R., и др. [10] (2019) предложили CNN, модифицированную на основе LeNet, для определения четырех категорий кукурузы (три класса болезней и один класс здоровья) с точностью 97,89%. Предложенная модель больше подходит для выявления болезней листьев кукурузы на сложном фоне. На основе базы данных PlantVillage для кукурузы изучается процесс построения моделей для мультиклассовой задачи определения болезней кукурузы

Методы исследования

В этой статье набор данных разделен на набор данных для обучения и набор данных для проверки в соответствии с принципами 3 к 1 из-за баланса выборки. Для классификации кукурузы по четырем категориям (рис.2), а именно: здоровая (H), южная пятнистость листьев кукурузы (SCLB) (Aregbesola et al., 2020). Набор данных изображений, который включал 7316 изображений, состоит из двух частей, а именно, одна часть собрана из общедоступного набора данных Plant Village. Для реализации данной задачи мы загружаем необходимый пакет программ и базу данных болезней растений. Рассмотрим подробно процесс построения модели распознавания болезней кукурузы.

```
import pandas as pd
import cv2 as cv
import numpy as np
import matplotlib.pyplot as plt
import os
import seaborn as sns
```

```
DATASET="D:/Dataset_Plant_Village_full/train_corn"
DATASET2="D:/Dataset_Plant_Village_full/valid_corn"
```

Графическое представление данных болезней кукурузы для обучения и проверки представлены на Рис.1. и Рис.2. соответственно.

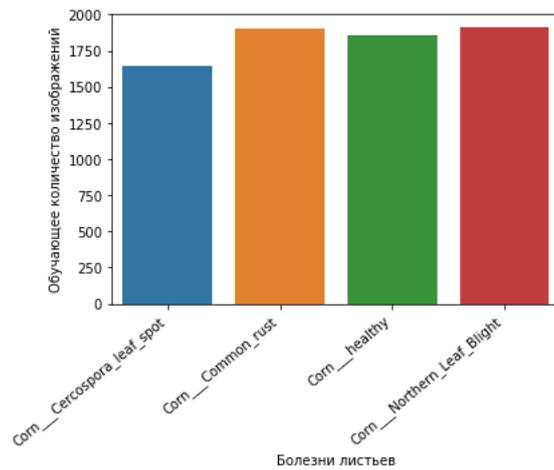


Рис.1. Количество обучающего набора данных для обучения модели.

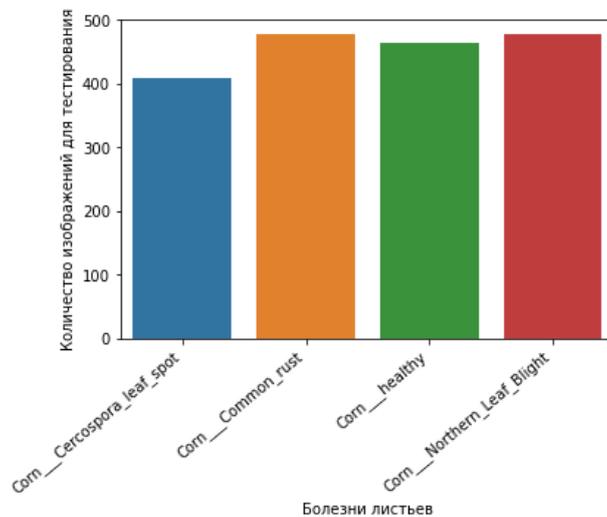


Рис.2. Количество проверочного набора данных для обучения модели .

Модель обучается на категориальных данных, следующими данными параметрами

```
hist=classifier.fit(X_train,one_hot_train,epochs=75,batch_size=128,validation_split=0.2)
```

где one_hot_train=to_categorical(y_train) соответствует болезням растения следующим образом.

```
Y=[]
```

```
for i in y_train:
```

```
    if i==0:
```

```
        Y.append("Corn__Cercospora_leaf_spot")
```

```

elif i==1:
    Y.append("Corn___Common_rust")
elif i==2:
    Y.append("Corn___healthy")
else:
    Y.append("Corn___Northern_Leaf_Blight")

```

```

classifier.summary()
Model: "sequential_2"

```

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 62, 62, 32)	896
max_pooling2d_6 (MaxPooling 2D)	(None, 31, 31, 32)	0
dropout_6 (Dropout)	(None, 31, 31, 32)	0
conv2d_7 (Conv2D)	(None, 29, 29, 64)	18496
max_pooling2d_7 (MaxPooling 2D)	(None, 14, 14, 64)	0
dropout_7 (Dropout)	(None, 14, 14, 64)	0
conv2d_8 (Conv2D)	(None, 12, 12, 128)	73856
max_pooling2d_8 (MaxPooling 2D)	(None, 6, 6, 128)	0
dropout_8 (Dropout)	(None, 6, 6, 128)	0
flatten_2 (Flatten)	(None, 4608)	0
dense_8 (Dense)	(None, 64)	294976
dense_9 (Dense)	(None, 128)	8320
dense_10 (Dense)	(None, 64)	8256
dense_11 (Dense)	(None, 4)	260

```

=====
Total params: 405,060
Trainable params: 405,060
Non-trainable params: 0

```

Обучение модели и ее точность

Epoch 1/75

46/46 [=====] - 11s 228ms/step - loss: 1.0202 - accuracy: 0.5236 - val_loss: 0.5931 - val_accuracy: 0.7439

.....
.....
Epoch 75/75

46/46 [=====] - 10s 224ms/step - loss:
0.0376 - accuracy: 0.9875 - val_loss: 0.1007 - val_accuracy: 0.9652

Точность модели, обученная на тестовых данных X_test и на категориальных данных one_hot_test

```
test_loss,test_acc=classifier.evaluate(X_test,one_hot_test)  
test_loss,test_acc
```

```
58/58 [=====] - 1s 12ms/step - loss: 0.1414 - accuracy: 0.9661  
(0.14138096570968628, 0.9661017060279846)
```

В графическом представлении процесса обучения модели и ошибка модели на данных обучения и проверки составляет 14.13%. Модель почти не переобучена.

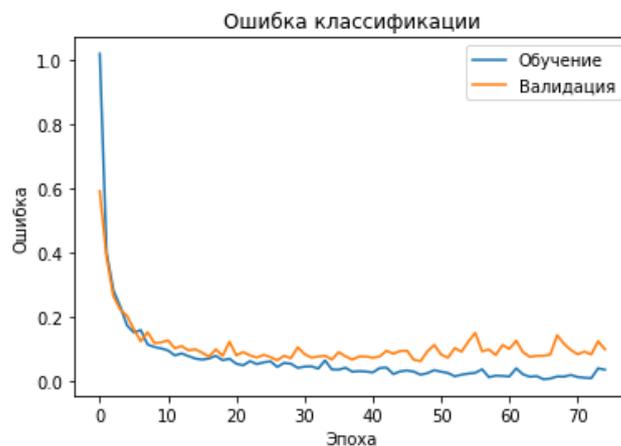


Рис.3. Ошибка модели на обучающем и тестовом множестве данных
Точность модели на обучающих и проверочных данных достигает до 0.9652 т.е. 96.52%

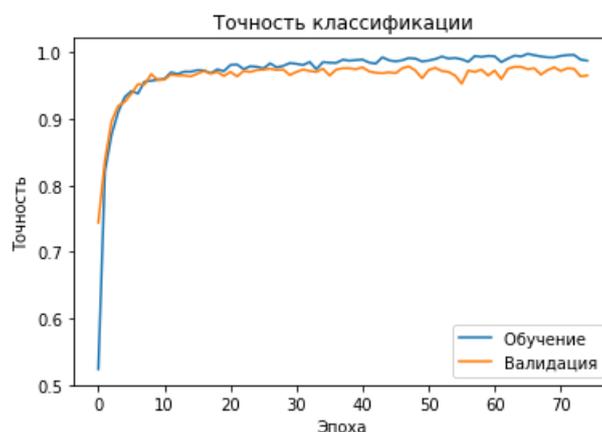


Рис.4. Ошибка модели на обучающем и тестовом множестве данных
 Рассмотрим теперь кривую ROC, которая обозначает кривую рабочих характеристик созданной модели. На рис.5 графическое представление эффективности построенной модели для задачи классификации болезни кукурузы. На графике пороги классификации задачи.

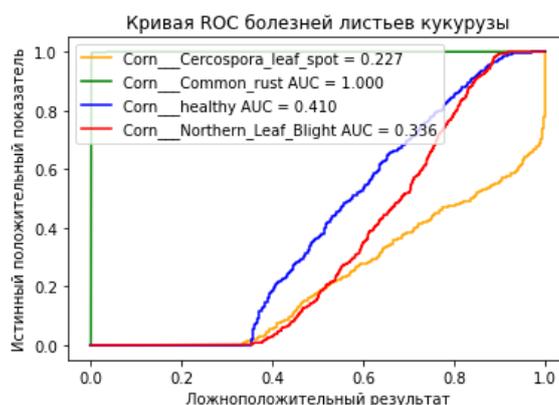


Рис.5. Кривая ROC для болезней кукурузы.

Теперь изучим матрицу погрешностей. Вот полученные результаты

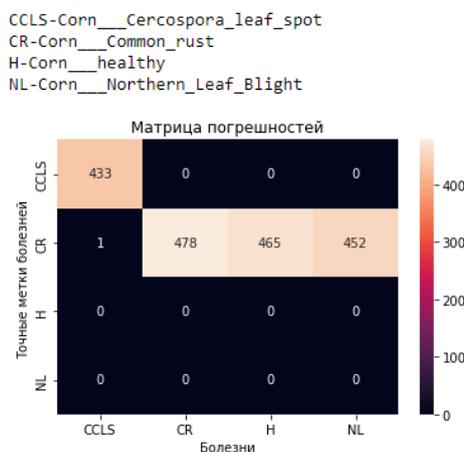


Рис.6. Матрица погрешностей для болезней кукурузы.

Матрица ошибок — это комбинация нашего прогноза (1 или 0) и фактического значения (1 или 0). В зависимости от результата предсказания и того, корректна ли была проведена классификация, матрица разделена на 4 части. Например, true positive (истинно положительный) результат — это количество случаев, в которых мы правильно классифицируем модель как положительный. А false positive (ложноположительный) — это число случаев, в которых мы ошибочно классифицируем модель как положительный. Как видно из рисунка CCLS -Corn_Cercospora_leaf_spot правильно классифицировано в 433 изображениях, CR-Corn_Common_rust классифицировано в 478 изображениях правильно, H-Corn_healthy правильно классифицировано в 465 изображений и болезнь NL-Corn___Northern_Leaf_Blight в 452 случаях классифицировано правильно. Мы выше по ROC кривой установили, что 100% точная классификация модель установило болезнь Corn___Common_rust. Остальные болезни растений наш построенный модель установила с различными процентами точности.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Заключение. В данной статье построено модель распознавания болезни сельскохозяйственных культур. В частности построены модели распознавания болезни кукурузы, которая считается стратегической продукцией в сельскохозяйственной отрасли. Изучено процесс выявления трех видов болезней данной культуры. Проверено эффективность моделей с помощью ROC кривых. Даны также, оценки

погрешностей прогнозирования в виде матрицы погрешностей. Построенные модели основаны на применении нейронных технологий.

ЛИТЕРАТУРА

1. Ranum, P., Peña-Rosas, J. P., and Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Ann. N.Y. Acad. Sci.* 1312, 105–112. doi: 10.1111/nyas.12396
2. PubMed Abstract | CrossRef Full Text | Google Scholar
3. Savary, S., Ficke, A., Aubertot, J.-N., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security* 4, 519–537. doi: 10.1007/s12571-012-0200-5. CrossRef Full Text | Google Scholar
4. Ouppaphan, P. (2017). “Corn disease identification from leaf images using convolutional neural networks,” in *Proceedings of the 2017 21st International Computer Science and Engineering Conference (ICSEC)*, (Piscataway, NJ: IEEE), 1–5. doi: 10.1109/ICSEC.2017.8443919. CrossRef Full Text | Google Scholar
5. Saleem, M. H., Potgieter, J., and Mahmood Arif, K. (2019). Plant disease detection and classification by deep learning. *Plants* 8:468. doi: 10.3390/plants8110468
6. PubMed Abstract | CrossRef Full Text | Google Scholar
7. Panigrahi, K. P., Das, H., Sahoo, A. K., and Moharana, S. C. (2020). “Maize leaf disease detection and classification using machine learning algorithms,” in *Progress in Computing, Analytics and Networking*, eds P. K. Pattnaik, S. S. Rautaray, H. Das, and J. Nayak (Cham: Springer), 659–669. doi: 10.1155/2022/6504616. PubMed Abstract | CrossRef Full Text | Google Scholar
8. Liu, J., and Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review. *Plant Methods* 17, 1–18. doi: 10.1186/s13007-021-00722-9. PubMed Abstract | CrossRef Full Text | Google Scholar

9. Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Adv. Neural Inform. Process. Syst.* 25, 1097–1105. [Google Scholar](#)
1. Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Kavita, Ijaz, M. F., et al. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors* 21:4749. doi: 10.3390/s21144749. [PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
10. DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E. L., Yosinski, J., Gore, M. A., et al. (2017). Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning. *Phytopathology* 107, 1426–1432. doi: 10.1094/PHYTO-11-16-0417-R [PubMed Abstract](#) | [CrossRef Full Text](#) | [Google Scholar](#)
11. Ahila Priyadharshini, R., Arivazhagan, S., Arun, M., and Mirnalini, A. (2019). Maize leaf disease classification using deep convolutional neural networks. *Neural Comput. Applic.* 31, 8887–8895. doi: 10.1007/s00521-019-04228-3. [CrossRef Full Text](#) | [Google Scholar](#)