

УДК 004.056

ИДЕНТИФИКАЦИЯ БОЛЕЗНЕЙ РАСТЕНИЙ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

Сабитов¹Б.Р., Сейткадиева² Н.С., Осмонов³ Э.Т., Калтаев³ Б.Э

¹Кыргызский национальный университет им.Ж.Баласагына

²Кыргызский государственный университет им.И.Арабаева

³Кыргызский государственный университет строительства, транспорта и архитектуры им.Н.Исанова

Для решения проблемы распознавания болезней кукурузы в режиме реального времени в данной работе была предложена нейронная сеть, которую можно обобщить для построения модели идентификации болезней многих сельскохозяйственных культур. При оптимизации модели используется алгоритм Адама для нейронных сетей. В целях уменьшения переобученных моделей была использована регуляризация L2, и выходная информация моделируется с помощью нелинейной функции ReLU. Ядра входных изображений свертки имеют размерность 3x3. Отношение обучающей выборки к контрольной выборке составляет 3:1. При экспериментальном сравнении точность предложенной в данной статье модели распознавания болезней кукурузы в наборе данных составляет 93,6%. Изучено влияние метода увеличения данных на точность модели глубокого обучения.

Ключевые слова. Идентификация, нейронные сети, кукуруза, нейронная сеть, сверточные нейронные сети обучающая выборка, тестовая выборка.

ӨСҮМДҮК ООРУЛАРЫН НЕЙРЛҮК ТАРМАКТАР МЕНЕН АНЫКТОО

Сабитов¹Б.Р., Сейткадиева²Н.С., Осмонов³Э.Т., Калтаев³Б.Э.

¹Ж.Баласагын атындагы Кыргыз улуттук университети

²И.Арабаев атындагы Кыргыз мамлекеттик университети

³Н.Исанов атындагы Кыргыз мамлекеттик курулуш, транспорт жана архитектура университети

Жүгөрүдөгү ооруларды реалдуу убакыт режиминде таануу маселесин чечүү үчүн бул иште көптөгөн айыл чарба өсүмдүктөрүнүн ооруларын аныктоонун моделин жалпылаштырууга мүмкүн болгон нейрондук түйүн сунушталган. Моделди оптималдаштырууда Адамдын нейрондук түйүндөр үчүн алгоритми колдонулат. Ашыкча моделдерди азайтуу үчүн L2 регуляризациясы колдонулган жана чыгаруу сызыктуу эмес ReLU функциясы менен моделдештирилген. Киргизилген конволюция сүрөттөрүнүн өзөктөрү 3x3 өлчөмүнө ээ. Тиешелүүлүк үлгүсүнүн контролдук үлгүгө катышы 3:1. Эксперименталдык салыштырганда, жүгөрү оорусун таануу моделинин тактыгы маалыматтарда ушул макалада сунушталган. белгиленген 93,6% түзөт. Терең үйрөнүү моделинин тактыгына маалыматтарды көбөйтүү ыкмасынын таасири изилденген.

Баштапкы сөздөр. Идентификация, нейрон тармактары, жүгөрү, нейрондук тармак, конволюциялык нейрон тармактары боюнча тренингдер топтому, тест топтому.

IDENTIFICATION OF PLANT DISEASES USING NEURAL NETWORKS

Sabitov¹B.R., Seitkazieva²N.S., Osmonov³E.T., Kaltaev³B.E

¹Kyrgyz National University named after Zh. Balasagyn

²Kyrgyz State University named after I. Arabaev

³Kyrgyz State University of Construction, Transport and Architecture
named of N.Isanov

To solve the problem of real-time recognition of diseases in corn, a neural network was proposed in this work, which can be generalized to build a model for identifying diseases of many crops. When optimizing the model, Adam's algorithm for neural networks is used. In order to reduce overfitting models, L2 regularization was used and the output is modeled with a non-linear ReLU function. The kernels of the input convolution images have a dimension of 3 x 3. The ratio of the training sample to the control sample is 3: 1. In an experimental comparison, the accuracy of the corn disease recognition model proposed in this article in the data set is 93.6%. The influence of the data augmentation method on the accuracy of a deep learning model was studied.

Keywords. Identification, neural networks, corn, neural network, convolutional neural networks training set, test set.

Введение

Кукуруза является важной продовольственной, кормовой и промышленной культурой для многих районов Кыргызской Республики. Это

культура как и картофель, люцерна и ячмень имеет стратегическое значение в сельском хозяйстве. Из нее многие фермерские хозяйства заготавливают корма для содержания животных и птиц. Площадь посевов кукурузы и общий объем производства кукурузы в КР в связи с нарастающей потребностью растет с каждым годом. С развитием производства кукурузы появилось и много видов болезней кукурузы, большинство из которых вызывается грибами, бактериями и вирусами [1]. Как быстро и точно диагностировать болезни кукурузы и принимать соответствующие меры контроля, имеет большое значение для производства кукурузы. При визуальных наблюдениях фермеры, основываясь на многолетнем опыте ставят диагноз для лечения и в большинстве случаев получается неправильные диагнозы.

Данный метод отнимает много времени, трудоемко и расходует, а заболевание кукурузы невозможно вовремя диагностировать и лечить, что приводит к низкой эффективности производства кукурузы. В последнее время с развитием технологий компьютерного зрения и использованием методов распознавания изображений, стало важным направлением исследований в области диагностики и обнаружения многих болезней сельскохозяйственных культур [2]. Сверточные нейронные сети (CNN) для распознавание болезней растений на основе изображений в основном основано на традиционных методах обработки изображений и методах, Yang et al. [3], Чен и др. [4], Чжан С. и Чжан С. [5], а также Чжу и Сюэ [6] использовали метод классификации SVM (машина опорных векторов), генетический алгоритм, локальное дискриминантное картирование (LDP) и алгоритм локального линейного встраивания (LLE) на стороне сервера, чтобы уменьшить извлеченные признаки болезни, с правильным уровнем распознавания 93,2%, 90,0%, 94,4% и 99,5% соответственно.

Вышеупомянутое исследование методов распознавания заболеваний, основанных на традиционной технологии обработки изображений, дало некоторые результаты, но работа этих методов слишком громоздка,

надежность низкая, а метод выделения признаков не является универсальным, что делает способность к обобщению всего метод плохой.

В последние годы метод глубокого обучения также широко используется для выявления заболеваний кукурузы. Сбор данных больных и здоровых растений представляет собой трудоемкую и затратную работу. В связи с этим в данной работе используется открытые данные для обучения модели с применением специально построенной нейронной сети. Для тестирования модели используется данные собранные вручную по 4 классам.

Объектами исследования являются мозаичная болезнь кукурузы, серая пятнистость, ржавчина и пятнистость листьев. Изображения болезней кукурузы классифицируются и идентифицируются, а также подготавливаются изображения здоровой кукурузы. Всего было использовано 7316 изображений четырех классов кукурузы, три вида болезней и здоровые листья кукурузы, а также 1829 изображений использованы для валидации модели. При использовании нейронных технологий, обычно требуется большое количество изображений.

Для решения данной проблемы использовано метод увеличения данных, после применения которой, можно добиться значительного увеличения количества изображений в наборе данных. Мы выполняем различные операции над изображениями, такие как сдвиг, поворот, масштабирование и зеркальное отображение набора данных, чтобы разнообразить наш набор данных.

Предоставление модели дополненных изображений помогает ей эффективно изучать особенности разных областей одного изображения и, таким образом, лучше работать с невидимыми данными изображения. Используется класс `flow_directory` при использовании генератора `train_datagen`. Методы глубокого обучения добились хороших результатов в исследованиях выявления болезней кукурузы, особенно на основе модели ResNet. С применением модели ResNet не только можно добиться

повышения точности распознавания, но и повысить скорость распознавания. Предварительный просмотр batch изображений-случайный первоначальный ввод изображений в нейронную сеть выглядит следующим образом:

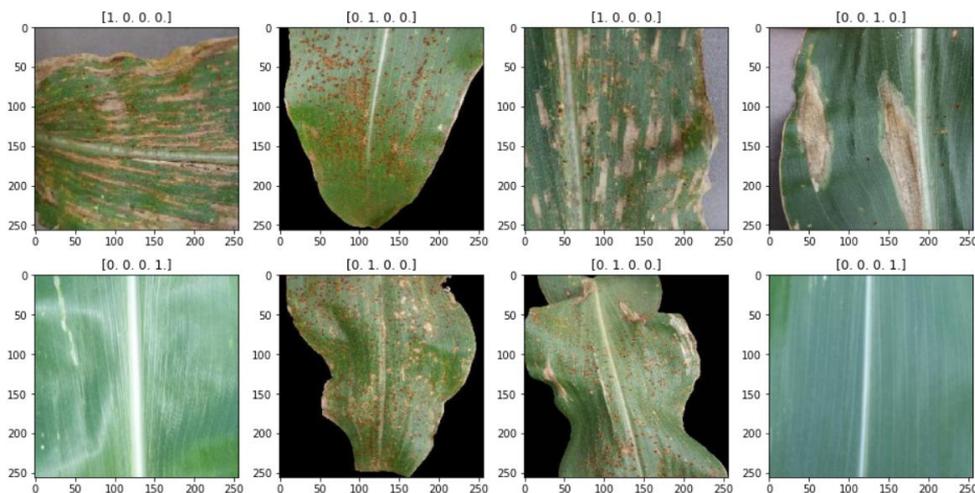


Рис.1. Первые 8 изображений из batch_size=64.

Чтобы, повысить точность идентификации болезней кукурузы, для расширения набора данных использовались различные методы улучшения данных, такие как осветление, перевод и отражение. Количество изображений мозаичной болезни в исходном наборе данных увеличилось в четыре раза, количество изображений серых пятен увеличилось с 2,5 раза, а количество изображений ржавчины увеличилось почти 4,5 раза. Количество изображений пятнистости листьев увеличилось до 4 раза. Было увеличено и количество обычных здоровых изображений кукурузы. Пример изображения из расширенного пакета данных показан на рисунке 2.

```
plt.imshow( x_batch[12] );  
plt.grid(False)  
plt.show();
```

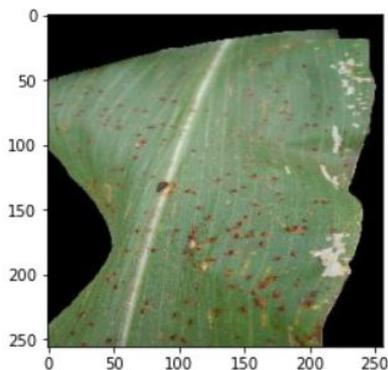


Рис.2. Случайный просмотр изображений с меткой layers =13

Данные до увеличения данных

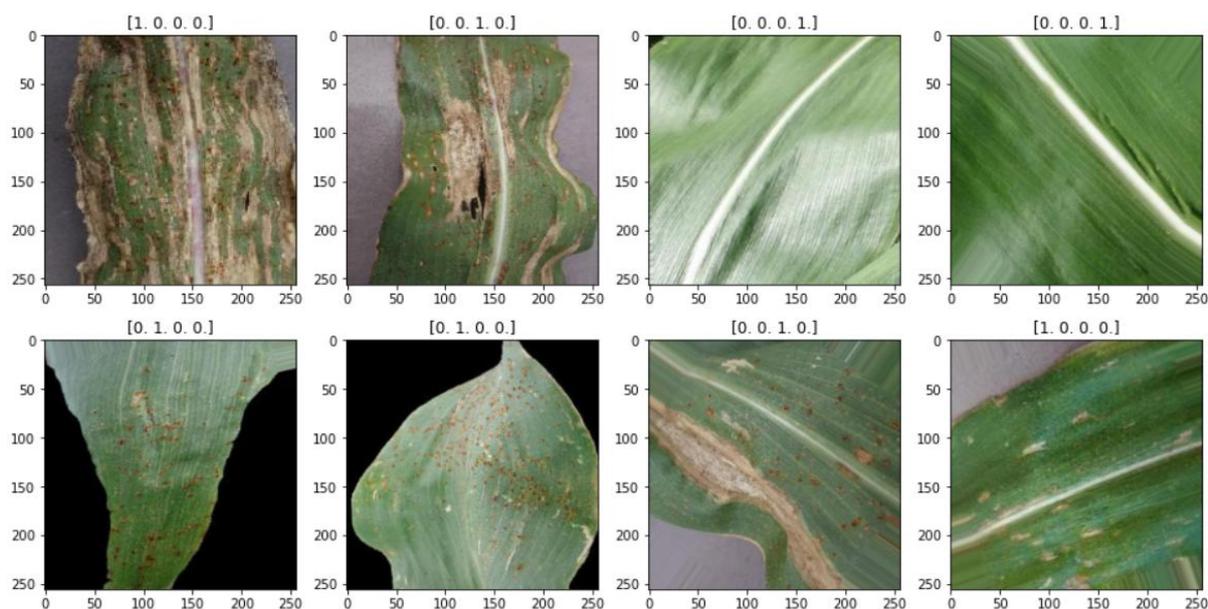


Рис.3.Изображения из пакета до применения генератора увеличения данных

Теперь рассмотрим объединение нескольких методов увеличения данных. Комбинированное увеличение данных и отбеливание. Отбеливание означает уменьшение избыточного изображения до требуемого масштаба и называется ZCA. Отбеливающее преобразование, таким образом минимизирует избыточность в матрице пиксельных изображений. Интуитивно он выделяет пиксели с высокой дисперсией на изображения с низкой дисперсией. При отбеливании изображения используется тот же метод уменьшения размеров, что и в анализе основных компонентов машинного обучения PCA, альтернатива которого называется ZCA. Однако преобразованное изображение сохраняет те же размеры исходного изображения. Для выполнения преобразование отбеливания ZCA, необходимо установить аргумент `zca_whitening` в значение `True` в `ImageDataGenerator`. Вот код примера отбеливания и увеличения данных:

```

augmented_zca_train_datagen = ImageDataGenerator(rotation_range=90,
width_shift_range=0.2,
height_shift_range=0.2,
zoom_range=0.4,
zca_whitening=True)

augmented_zca_train_generator = augmented_zca_train_datagen.flow_from_directory(train_image_dir,
target_size=(256, 256),
batch_size=128,
class_mode='categorical')

```

Данные после увеличения и отбеливания данных, представлен в виде таблицы изображений:

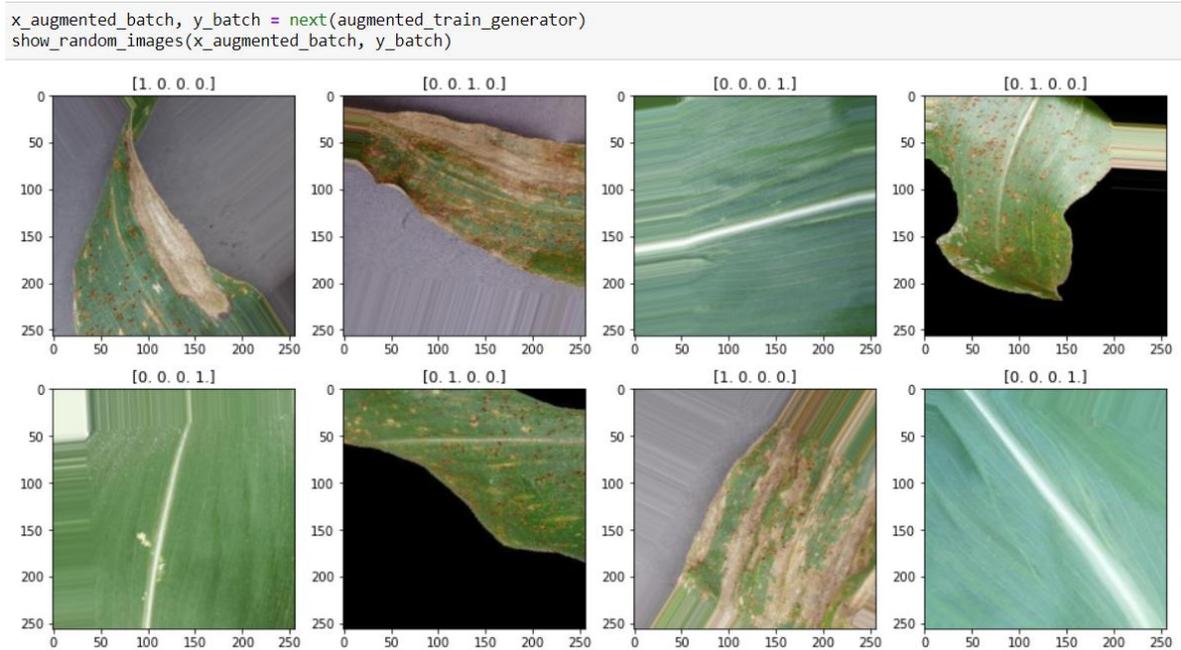


Рис.4.Изображения полученная после отбеливания границ изображений.

Теперь построим сверточную нейронную сеть с входными изображениями определяемые с помощью четырёх классов и индексацией меток болезней кукурузы по их типу

```

train_generator.class_indices
{'Corn__Cercospora_leaf_spot Gray_leaf_spot': 0,
 'Corn__Common_rust': 1,
 'Corn__Northern_Leaf_Blight': 2,
 'Corn__healthy': 3}

image_shape = train_generator.image_shape

train_generator.n

```

Создадим нейронная сеть, которая использует нелинейную функцию

```

model = Sequential()
model.add(Conv2D(filters=16, kernel_size=(3,3), strides=1, padding='same', input_shape=image_shape))
model.add(LeakyReLU(0.1))

model.add(MaxPooling2D(pool_size=(3, 3)))

model.add(Conv2D(filters=32, kernel_size=(3,3), strides=1, padding='same', input_shape=image_shape))
model.add(LeakyReLU(0.1))

model.add(MaxPooling2D(pool_size=(3, 3)))

model.add(Conv2D(filters=64, kernel_size=(3,3), strides=1, padding='same', input_shape=image_shape))
model.add(LeakyReLU(0.1))

model.add(MaxPooling2D(pool_size=(3, 3)))

model.add(Flatten())

model.add(Dense(256))
model.add(LeakyReLU(0.1))
model.add(Dropout(0.5))

model.add(Dense(64))
model.add(LeakyReLU(0.1))
model.add(Dropout(0.5))

model.add(Dense(4))
model.add(Activation('softmax'))

```

Обучим сеть для построения модели:

```

%%time

history = model.fit_generator(augmented_train_generator,
                             steps_per_epoch=10,
                             epochs=50,
                             callbacks=callbacks_list,
                             validation_data=augmented_train_generator,
                             validation_steps=5)

```

Суть оптимизации структуры модели заключается в минимизации повторения функции потерь. В этой статье для оптимизации модели используется адаптивная оценка модели (Адам) вместо традиционного стохастического градиентного метода SGD (стохастический градиентный спуск). Суть алгоритма Адама заключается в динамической настройке оценки момента первого порядка (бета1) и оценки момента второго порядка (бета2) каждого градиента параметра в соответствии с функцией потерь. Он характеризуется высокой вычислительной эффективностью и небольшой памятью, что подходит для решения задач большой выборки и оптимизации параметров.

```

model_zca.compile(optimizer='adam',
                  loss = 'categorical_crossentropy',
                  metrics = ['accuracy'])

```

Сохраним модель для ее дальнейшего использования.

```
import pickle
```

```
pickle.dump( history.history, open('C:/Model/history_aug_nozca_Corn.pkl', 'wb') )
```

Ниже

представлены две обученные модели нейронной сети. Точность модели до увеличения данных

```
plot_train_val_accuracy( history.history )
```

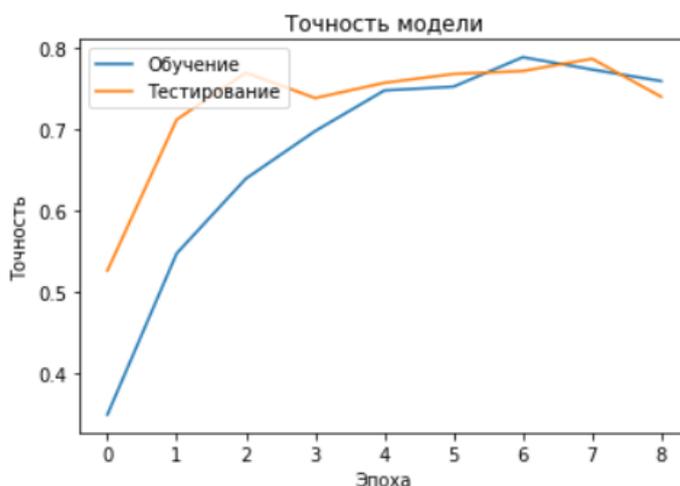


Рис.5.Точность модели до увеличения данных на обучающей и тестовой множестве данных.

Точность модели после обучения с увеличенными данными

```
plot_train_val_accuracy(history_zca.history)
```

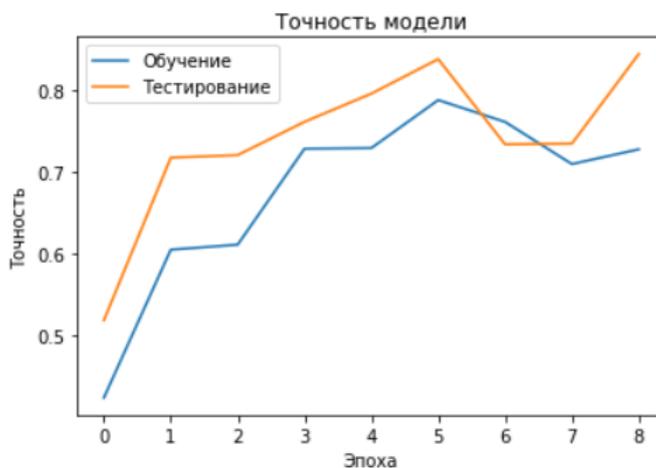


Рис.6.Точность модели после увеличения данных на обучающей и тестовой множестве данных.

Заключение. В данной статье, основанной на структуре модели ResNet50, метод экспоненциального затухания используется для настройки скорости обучения, а регулярный член L2 добавляется к функции кросс-энтропии, чтобы уменьшить вес. Во избежание чрезмерного обучения в процессе обучения между слоями сети используются стратегия отсева Dropout и нелинейная функция стимулирования ReLU. Первый слой модели ResNet50 был изменен на три сверточных слоя 3x3 для повышения точности распознавания небольших очагов болезней кукурузы.

Изучена распознавания трех болезней кукурузы, в том числе мозаичная болезнь кукурузы, болезнь серых листьев и пятнистость листьев. Уровень распознавания обучающего набора составляет 93,52%, а средний правильный показатель набора проверки без обучения составляет 91,826%, что в целом обеспечивает удовлетворительный эффект распознавания.

ЛИТЕРАТУРА

1. Х. Ван, «Определение названий шести важных болезней и патогенов кукурузы», Китайская сельскохозяйственная наука, том. 51, нет. 18, стр. 3497–3507, 2018. Посмотреть в: [Академия Google](#)
2. Фу Л., Хе Р., Лю П. Новый инвариантный момент и система распознавания болезней кукурузы на основе нейронной сети // Компьютерная инженерия и приложения . 48, нет. 4, стр. 219–221, 2012. Посмотреть в: [Академия Google](#)
3. Б. Ян, Л. Чжан и С. Ян, «Исследование технологии выявления болезней кукурузы на основе мобильного Интернета и технологии SVM», Сельскохозяйственная наука Цзилинь , том. 39, нет. 3, стр. 59–62, 2014. Посмотреть в: [Академия Google](#)
4. Л. Чен и Л. Ван, «Применение вероятностной нейронной сети для распознавания болезней листьев кукурузы», Исследование механизации сельского хозяйства , том. 48, нет. 6, стр. 145–148, 2011. Посмотреть в: [Академия Google](#)

5. С. Чжан и К. Чжан, «Метод выявления болезней кукурузы на основе алгоритма локального дискриминантного картирования», Журнал сельскохозяйственной инженерии , том. 30, нет. 11, стр. 167–172, 2014.Посмотреть в: [Академия Google](#)
6. Дж. Чжу и Л. Сюэ, «Метод уменьшения размеров и кластеризации для распознавания цветных изображений болезни листьев кукурузы», Сельскохозяйственные науки Цзянсу , том. 44, нет. 7, стр. 350–354, 2016.Посмотреть в: [Академия Google](#)