КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ АЛЬ-ФАРАБИ

ЛЕКЦИЯ № 11

Верификация искусственного интеллекта при распознавании заболеваний растений агроробототом с машинным зрением

Бельгибаев Бауржан Абдрахимович
Д.т.н., доцент · Кафедра «Искусственный интеллект и big data»
«Технологии в интеллектуальных системах управления в IoT» · 4 курс, бакалавриат

1 Цель лекции

Сформировать системное представление о проверке корректности, обобщающей способности, устойчивости и пригодности к эксплуатации моделей компьютерного зрения, предназначенных для распознавания заболеваний растений в составе интеллектуального агроробота.

2 Задачи

- Разграничить обучение, валидацию, тестирование и верификацию модели.
- Изучить правила формирования независимого тестового набора изображений.
- Освоить матрицу ошибок и метрики качества классификации.
- Рассмотреть тестирование модели в изменяющихся полевых условиях.
- Разработать протокол испытаний системы распознавания на агророботе.

3 План лекции

1. Понятие верификации и жизненный цикл модели ИИ
2. Формирование тестовых наборов изображений растений
3. Метрики качества и математический аппарат
4. Проверка устойчивости и достоверности результатов
5. Особенности полевой эксплуатации модели на агророботе
6. Программная реализация проверки на Python
7. Экспериментальный протокол и интерпретация результатов
8. Практические ограничения и перспективы внедрения

1 Понятие верификации и жизненный цикл модели ИИ

Интеллектуальный агроробот использует изображения листьев и стеблей, полученные камерой, для оценки состояния растений. Модель машинного обучения преобразует входное изображение x в прогноз класса $\hat{y} = f\theta(x)$, где θ — обученные параметры. На выходе может формироваться вероятность наличия определённого заболевания либо класса «здоровое растение».

Обучение минимизирует функцию потерь на обучающих примерах. Валидация помогает выбрать архитектуру, гиперпараметры и порог решения. Независимое тестирование оценивает поведение уже выбранной модели на данных, не участвовавших в настройке. Верификация шире обычного тестирования: она устанавливает, удовлетворяет ли реализация заранее определённым требованиям и корректно ли выполнены вычисления и интеграция. Валидация применительно к реальной задаче дополнительно отвечает на вопрос, пригодна ли система для предполагаемого применения.

Следует различать корректность классификации изображений и корректность действий агроробота. Даже модель с высокой точностью может оказаться непригодной для управления опрыскиванием при больших задержках или высокой цене ложноположительного решения.

4 Контур верификации ИИ-системы



Рис. 20. Верификация повторяется при каждом изменении модели или условий эксплуатации.

2 Формирование тестовых наборов изображений растений

Качество теста зависит от независимости и репрезентативности выборки. Изображения одного растения, сделанные подряд, часто сильно коррелированы. Если часть таких кадров попадает в обучение, а часть — в тест, оценка может быть искусственно завышена. Для полевых исследований предпочтительно разделять данные по растениям, участкам, датам съёмки или хозяйствам, а не только по отдельным файлам.

Необходимо документировать культуру, сорт, заболевание, стадию поражения, условия освещения, тип камеры, расстояние до растения и способ подтверждения диагноза. Разметка должна быть проверена компетентным специалистом. Нельзя считать любую визуальную аномалию доказательством конкретного заболевания: дефицит питания, повреждение вредителями и засуха могут давать сходные признаки.

2 Формирование тестовых наборов изображений растений (продолжение)

Набор	Назначение	Основное требование
Обучающий	Настройка параметров	Достаточное разнообразие и корректная разметка
Валидационный	Выбор гиперпараметров и порога	Не использовать для финальной оценки
Тестовый	Независимая оценка качества	Исключить утечки и повторное использование при настройке
Внешний полевой	Оценка переноса на новые условия	Иные участки, даты, устройства или хозяйства

2 Формирование тестовых наборов изображений растений (продолжение)

Стратификация по классам помогает контролировать представительство заболеваний. Однако естественная распространённость заболеваний в поле может отличаться от баланса в тестовой выборке; это влияет на положительную прогностическую ценность. Поэтому целесообразно дополнительно оценивать модель на выборке с реалистичной частотой классов.

3 Метрики качества и математический аппарат

Для бинарного решения «заболевание обнаружено / не обнаружено» используются TP — верно выявленные случаи, TN — верно исключённые случаи, FP — ложные срабатывания, FN — пропущенные заболевания.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN).$$

$$Precision = TP / (TP + FP).$$

$$Recall\ (чувствительность) = TP / (TP + FN).$$

$$Specificity = TN / (TN + FP).$$

$$F1 = 2 \cdot Precision \cdot Recall / (Precision + Recall).$$

$$Balanced\ Accuracy = (Recall + Specificity) / 2.$$

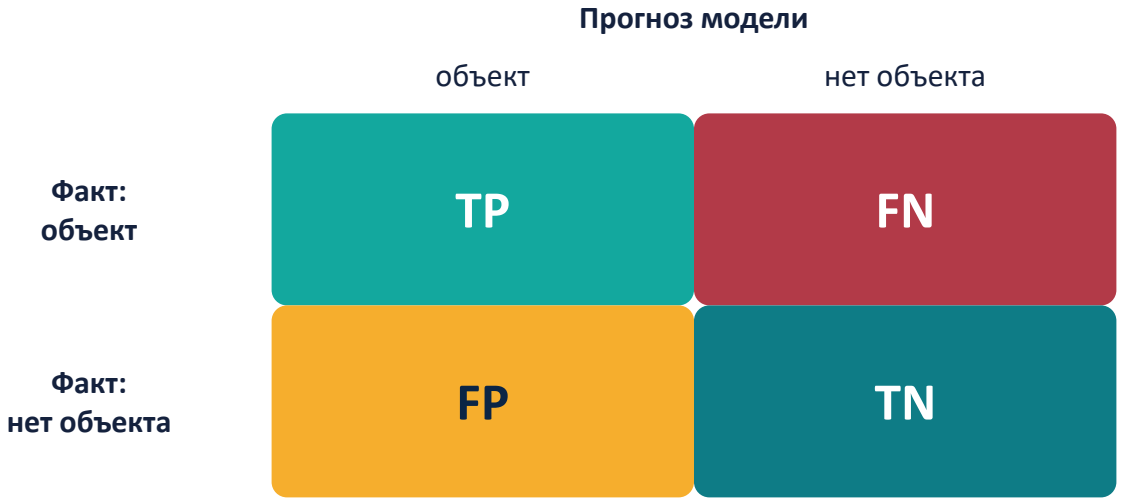
Метрики качества и математический аппарат (продолжение)

Пример (условные учебные данные): на 200 независимых изображениях получено TP = 72, FN = 8, FP = 12, TN = 108. Тогда Accuracy = 180/200 = 0,900; Precision = 72/84 ≈ 0,857; Recall = 72/80 = 0,900; Specificity = 108/120 = 0,900; F1 ≈ 0,878. Эти значения иллюстративны и не являются результатом реального эксперимента.

Для многоклассовой классификации формируется матрица Cij, где строка i соответствует истинному классу, а столбец j — предсказанному. По каждому классу рассчитываются Precision, Recall и F1. Macro-F1 усредняет F1 по классам без учёта частоты, weighted-F1 — с весами по числу примеров. При редких заболеваниях одной Accuracy недостаточно.

Порог вероятности τ задаёт правило $\hat{y} = 1$, если $P(y=1|x) \geq \tau$. Изменяя τ, получают различные компромиссы между ложными срабатываниями и пропусками. ROC-AUC и PR-AUC оценивают качество ранжирования по порогам; при сильном дисбалансе классов PR-AUC часто информативнее.

6 Матрица ошибок и метрики



Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
F1 = 2PR / (P + R)

Рис. 21. Метрики качества вычисляются из четырёх исходов классификации.

4 Проверка устойчивости и достоверности результатов

Полевые изображения подвержены изменениям освещённости, теням, размытию, перекрытиям листьев, погодным условиям и сдвигу распределения данных. Поэтому необходимо отдельно оценивать качество на заранее заданных подгруппах: яркий свет, пасмурная погода, частичное перекрытие, разные расстояния, различные сорта и стадии поражения.

Для оценки неопределённости можно использовать доверительные интервалы для долей, бутстрэп по независимым растениям или участкам, а также повторные испытания на нескольких сезонах. Важно выбирать единицу статистической независимости: сотни кадров одного растения не равнозначны сотням независимых растений.

Калибровка вероятностей показывает, насколько частота верных решений соответствует заявленной уверенности модели. Например, среди прогнозов с уверенностью около 0,8 корректными должны быть приблизительно 80% при хорошей калибровке. Для её оценки используют reliability diagram и Brier score: $BS = (1/N)\ \Sigma(p_i - y_i)^2$.

5 Особенности полевой эксплуатации модели на агророботе

Типовая цепочка обработки: камера → захват кадра → предварительная обработка → нейросетевая модель → фильтрация и объединение решений → модуль управления → исполнительный механизм или уведомление оператору. Система должна обеспечивать прослеживаемость решений, регистрацию ошибок и безопасное поведение при потере связи или низкой уверенности прогноза.

Время цикла складывается из задержек получения кадра, предварительной обработки, инференса, передачи и формирования команды: $T_{цикл} = T_{кам} + T_{пред} + T_{ИИ} + T_{сеть} + T_{упр}$. При частоте камеры f и времени цикла $T_{цикл}$ условие $T_{цикл} \le 1/f$ желательно для обработки каждого кадра без накопления очереди; при нарушении необходимо пропускать кадры или снижать частоту.

Если робот движется со скоростью v, то за время задержки $T_{цикл}$ он проходит расстояние $\Delta s = v \cdot T_{цикл}$. При $v = 0{,}5$ м/с и $T_{цикл} = 0{,}3$ с $\Delta s = 0{,}15$ м. Такая ошибка пространственной привязки может быть существенной при точечной обработке растений. Требуются синхронизация временных меток, учёт положения камеры и исполнительного механизма, а также безопасные ограничения.

6 Программная реализация проверки на Python

Ниже приведён минимальный пример оценки уже полученных истинных и предсказанных меток. В реальном исследовании метки должны быть получены из независимого тестового набора, а версия модели и настройки — зафиксированы до запуска.

```
from sklearn.metrics import (confusion_matrix, classification_report,
                             accuracy_score, balanced_accuracy_score)
y_true = [1]*72 + [1]*8 + [0]*12 + [0]*108
y_pred = [1]*72 + [0]*8 + [1]*12 + [0]*108
print(confusion_matrix(y_true, y_pred, labels=[0, 1]))
print(classification_report(y_true, y_pred,
                            target_names=["healthy", "disease"],
                            zero_division=0))
print("Accuracy:", accuracy_score(y_true, y_pred))
print("Balanced accuracy:", balanced_accuracy_score(y_true, y_pred))
```

Для работы с реальными изображениями следует использовать неизменяемый список файлов тестовой выборки, фиксировать контрольные суммы, версии библиотек и весов модели. Результаты удобно сохранять в CSV: идентификатор изображения, истинный класс, предсказанный класс, вероятность, время обработки, параметры съёмки.

7 Экспериментальный протокол и интерпретация результатов

1. Сформулировать назначение модели и требования к чувствительности, специфичности и задержке.
2. Зафиксировать архитектуру, веса, порог решения и программную среду до испытания.
3. Сформировать независимую тестовую выборку по растениям и участкам.
4. Проверить качество разметки и состав классов.
5. Выполнить пакетное тестирование, построить матрицу ошибок и рассчитать метрики.
6. Провести стратифицированный анализ по условиям съёмки и заболеваниям.
7. Измерить задержки, энергопотребление и пропускную способность на целевом оборудовании.
8. Провести ограниченные полевые испытания с контролем оператором, документировать отказы.
9. Сопоставить результаты с критериями допуска и описать ограничения переноса.

Для сравнения двух моделей на одинаковых объектах полезны парные процедуры оценки. Не следует делать вывод о превосходстве одной модели только по разнице процентов без анализа доверительных интервалов, размера выборки и распределения ошибок.

8 Практические ограничения и перспективы внедрения

Система распознавания болезней растений должна рассматриваться как компонент поддержки решений. Автоматическое назначение агрохимической обработки только по результату одной нейросети может быть ошибочным. В производственной системе целесообразны подтверждение сомнительных случаев, ограничения по зонам применения и контроль со стороны агронома.

Перспективные направления исследований включают обучение на малом числе размеченных изображений, перенос обучения между культурами, обнаружение неизвестных классов, мультимодальное объединение RGB и спектральных данных, оценку неопределённости и адаптацию модели к изменениям сезона. При этом обновление модели требует повторной независимой проверки.

12 Заключение

Верификация модели компьютерного зрения для агроробота — это последовательный процесс проверки данных, алгоритма, вычислительной платформы и поведения всей системы управления. Надёжность определяется не только средней точностью классификации, но и чувствительностью к редким заболеваниям, устойчивостью к условиям съёмки, временем отклика и безопасностью исполнительных действий. Научно обоснованный протокол испытаний позволяет выявлять ограничения модели до внедрения в реальные агротехнологические процессы.

13 Контрольные вопросы

1. Чем различаются обучение, валидация, тестирование и верификация ИИ?
2. Почему нельзя разделять похожие кадры одного растения между обучением и тестом?
3. Как строится матрица ошибок бинарного классификатора?
4. В чём отличие Precision от Recall?
5. Почему Accuracy может вводить в заблуждение при редких заболеваниях?
6. Как рассчитывается F1-score?
7. Что означает изменение порога классификации?
8. Чем полезны ROC-AUC и PR-AUC?
9. Что такое калибровка вероятностей модели?
10. Какие факторы вызывают сдвиг распределения в поле?
11. Почему важно проводить испытания на независимых хозяйствах?

13 Контрольные вопросы (продолжение)

12. Как оценить суммарную задержку решения на агророботе?
13. Как задержка влияет на точность локальной обработки растения?
14. Какие данные необходимо сохранять в протоколе испытаний?
15. Какие ограничения безопасности следует предусмотреть при внедрении модели?

14 Рекомендуемая литература и ресурсы

- Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press, 2016.
- Bishop C. M. Pattern Recognition and Machine Learning. Springer, 2006.
- **Szeliski R. Computer Vision:** Algorithms and Applications. 2nd ed. Springer, 2022.
- **scikit-learn:** Model evaluation — https://scikit-learn.org/stable/modules/model_evaluation.html
- **OpenCV Documentation —** https://docs.opencv.org/
- **TensorFlow:** Image classification — https://www.tensorflow.org/tutorials/images/classification