Идентификация и подсчёт крупного рогатого скота при однокамерном и многокамерном видеонаблюдении

А. В. Лужецкая*, Д. С. Есаулова‡, К. Э. Караваева§, С. Н. Балавнев¶

Ручной учёт крупного рогатого скота трудоёмок; автоматический осложняется перекрытиями и повторными появлениями животных. Представлена система, объединяющая YOLO26L, ByteTrack и ResNet50. Для однокамерного подсчёта используются отбор изображений и объединение траекторий, для межкамерного сопоставления разработана геометрическая проверка. На MmCows получены mAP@0.5 $\approx$ 0,994 и Rank-1 Accuracy 90,6%.

Ключевые слова: компьютерное зрение, индивидуальная идентификация животных, многокамерное наблюдение, подсчёт животных.

1. Постановка задачи и метод

Автоматизация визуального учёта крупного рогатого скота направлена на снижение трудозатрат ручного подсчёта. Основной вклад работы состоит в разработке схемы накопительного однокамерного подсчёта с отбором изображений по качеству, положению и направлению животного для сопоставления траекторий и предотвращения повторного учёта особей.

Пусть S обозначает множество траекторий, полученных за интервал $[0, T]$. Требуется построить разбиение G множества S, в котором каждая группа содержит траектории одной и той же особи, а разные группы

* *Лужецкая Анастасия Викторовна* — аспирант факультета ВМК МГУ имени М. В. Ломоносова, e-mail: nastyajetskaya@yandex.ru.

Luzhetskaia Anastasiia Viktorovna — PhD student, Faculty of Computational Mathematics and Cybernetics, Lomonosov MSU.

‡ *Есаулова Диана Сергеевна* — студент факультета ВМК МГУ имени М. В. Ломоносова, e-mail: esaulovadiana2005@gmail.com.

Esaulova Diana Sergeevna — Student, Faculty of Computational Mathematics and Cybernetics, Lomonosov MSU.

§ *Караваева Ксения Эдуардовна* — студент физического факультета, МГУ имени М. В. Ломоносова, e-mail: karavaevake@my.msu.ru.

Karavaeva Ksenia Eduardovna — Student, Faculty of Physics, Lomonosov MSU.

¶ *Балавнев Семен Николаевич* — студент факультета психологии, МГУ имени М. В. Ломоносова, e-mail: semyon.balavnev@gmail.com.

Balavnev Semen Nikolaevich — Student, Faculty of Psychology, Lomonosov MSU.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

соответствуют разным особям. Оценка количества уникальных животных определяется числом групп: $\widehat{N}(T) = |G|$.Число особей заранее неизвестно. Для нескольких камер предполагаются синхронизация и известные матрицы проекции.

YOLO26L из Ultralytics [1], дообученная на MmCows [2], обнаруживает животных в заданной области загона; ByteTrack [3] формирует траектории. Изображения сохраняются для сопоставления после смены идентификатора трекера.

Из каждой траектории выбираются до 3 лучших изображений, затем 100 лучших суммарно. Оценка $q = \sum_{j=1}^{7} w_j q_j$ использует веса $w = (0{,}05; 0{,}25; 0{,}25; 0{,}15; 0{,}05; 0{,}18; 0{,}07)$. При $h(x) = \min(1, \max(0, x))$ компоненты равны:

$$q_1 = p; \quad q_2 = h(a/0{,}02); \quad q_3 = 1 - r; \quad q_4 = h(v/0{,}06);$$
$$q_5 = 1 - h(|m - 0{,}4|/0{,}5); \quad q_6 = 1 - h(e/0{,}15); \quad q_7 = 1 - h(u/0{,}30).$$

Здесь p обозначает уверенность детектора, a долю площади кадра, r долю рамки, перекрытую другими рамками, v дисперсию лапласиана, m среднюю яркость. Яркость нормирована к $[0, 1]$; e и u обозначают доли пикселей со всеми RGB-каналами $> 245/255$ и яркостью $< 0{,}05$ соответственно.

ResNet18 [4] различает классы «стоит/лежит»: 42 754 изображения, 80% для обучения со стратифицированным случайным разбиением, начальные веса ImageNet, 3 эпохи, перекрёстная энтропия, Adam с шагом 10^{-3} и размером батча 64. Сохраняются стоящие животные с вероятностью $\geq 0{,}5$. По маске BiRefNet [5] вычисляется $s = (L - R)/B$, где L, R обозначают среднюю толщину крайних 15% силуэта, B центральных 30%. При $s \geq 0{,}08$ голова считается направленной вправо; такие изображения отбираются.

ResNet50 [4] без классификатора формирует нормированные 2048-мерные признаки. Сходство траекторий равно максимальному косинусному сходству изображений. При значении $\geq 0{,}45$ траектории связываются; прямые и транзитивные связи образуют группы особей.

Для набора синхронизированных камер C с известной калибровкой центр рамки задаёт прямую $o_c + \lambda d_c$. Ближайшая к прямым точка определяется как

$$Q = \arg\min_{X \in \mathbb{R}^3} \sum_{c \in C} \left\| (I_3 - d_c d_c^\top)(X - o_c) \right\|_2^2.$$

Здесь o_c обозначает центр камеры, d_c единичное направление, I_3 единичную матрицу. Допустимы положительная глубина, среднее расстояние

до прямых $D \leq 90$ и максимальное попарное расстояние $P_{\max} \leq 120$ в единицах калибровки. Сначала рассматриваются кандидаты, прошедшие геометрическую проверку. В каждой из 2 групп кандидаты упорядочиваются по возрастанию оценки $0{,}7D + 0{,}3\overline{P}$, где $\overline{P}$ обозначает среднее попарное расстояние между прямыми. При одинаковой оценке предпочтение отдаётся большему визуальному сходству.

2. Экспериментальная проверка

Использованы данные 16 коров с 4 камер MmCows. YOLO26L обучалась 100 эпох на $\approx 10\,000$ изображений; на $\approx 6\,000$ валидационных получены mAP@0.5 $\approx 0{,}994$ и mAP@0.5:0.95 $\approx 0{,}845$.

Для ReID кропы каждой коровы отбирались каждые 2 мин с 4 камер за 24 ч. Первые 12 ч составили обучение и галерею, следующие 12 ч тест: по 23 040 изображений. ResNet50 дообучалась с весов ImageNet: TripletMarginLoss (отступ 0,3), MultiSimilarityMiner, Adam (10^{-4}), пакет 64, горизонтальные отражения, вариации яркости/контраста 0,2. Вход 256×128 нормировался по ImageNet. В этом эксперименте Rank-1 Accuracy составила 90,6%.

Корректность межкамерных сопоставлений определялась по идентификаторам животных в разметке MmCows. Среднее расстояние восстановленной точки до прямых наблюдения составило $\approx 8{,}7$ для корректных и $\approx 103{,}3$ для ошибочных пар. На видеозаписях проверена совместная работа модулей однокамерной системы: обнаружение животных, отбор изображений, повторная идентификация и объединение траекторий для подсчёта.

3. Заключение

Рассмотрены многокамерная схема с геометрической проверкой и однокамерная схема с отбором изображений и объединением траекторий. На данных MmCows получены оценки качества детекции и идентификации.

Список литературы

[1] Ultralytics, *Ultralytics YOLO26*, Documentation, accessed 09.10.2026, `https://docs.ultralytics.com/models/yolo26/`.

[2] H. Vu et al., “MmCows: A Multimodal Dataset for Dairy Cattle Monitoring”, *Advances in Neural Information Processing Systems*, **37**, 2024, 59451–59467. DOI: `10.52202/079017-1898`.

[3] Y. Zhang et al., “ByteTrack: Multi-object Tracking by Associating Every Detection Box”, *Computer Vision - ECCV 2022*, LNCS, **13682**, 2022, 1–21. DOI: 10.1007/978-3-031-20047-2_1.

[4] K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 770–778. DOI: 10.1109/CVPR.2016.90.

[5] P. Zheng et al., “Bilateral Reference for High-Resolution Dichotomous Image Segmentation”, *CAAI Artificial Intelligence Research*, **3** (2024), 9150038. DOI: 10.26599/AIR.2024.9150038 .

Статья поступила 9 октября 2026 г.

Cattle identification and counting in single-camera and multi-camera video monitoring

A. V. Luzhetskaia, D. S. Esaulova, K. E. Karavaeva, S. N. Balavnev

Manual cattle counting is labour-intensive, while automation is complicated by occlusions and repeated appearances. A system combining YOLO26L, ByteTrack and ResNet50 is presented. Image selection and track merging are used for single-camera counting, and geometric verification is developed for cross-camera matching. Experiments on MmCows yielded mAP@0.5 ≈ 0.994 and Rank-1 Accuracy of 90.6%.

Keywords: computer vision, individual animal identification, multi-camera monitoring, animal counting.

References

[1] Ultralytics, *Ultralytics YOLO26*, Documentation, accessed 09.10.2026, https://docs.ultralytics.com/models/yolo26/.

[2] H. Vu et al., “MmCows: A Multimodal Dataset for Dairy Cattle Monitoring”, *Advances in Neural Information Processing Systems*, **37**, 2024, 59451–59467. DOI: 10.52202/079017-1898.

[3] Y. Zhang et al., “ByteTrack: Multi-object Tracking by Associating Every Detection Box”, *Computer Vision - ECCV 2022*, LNCS, **13682**, 2022, 1–21. DOI: 10.1007/978-3-031-20047-2_1.

[4] K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 770–778. DOI: 10.1109/CVPR.2016.90.

[5] P. Zheng et al., “Bilateral Reference for High-Resolution Dichotomous Image Segmentation”, *CAAI Artificial Intelligence Research*, **3** (2024), 9150038. DOI: 10.26599/AIR.2024.9150038.

Received on October 9, 2026