

# Комплексный алгоритм анализа видеоданных колоноскопических исследований

Н. В. Котов

ЯрГУ им. П.Г. Демидова

n.kotov2@uniyar.ac.ru

**Аннотация.** Представлен комплексный алгоритм анализа видеоданных колоноскопических исследований для задачи контроля качества видеопоследовательностей с целью обнаружения области купола слепой кишки. Исследование проведено на собственной размеченной базе видеопоследовательностей колоноскопических исследований. Максимальная достигнутая точность составила 0,38 по метрике IoU для тестовых видеопоследовательностей.

**Ключевые слова:** колоноскопия, купол слепой кишки, анализ видеоизображений, оптический поток, глубокое обучение

## I. ВВЕДЕНИЕ

Процедура колоноскопии применяется для диагностики заболеваний толстой кишки, в том числе для выявления злокачественных образований. Важным критерием полноты исследования является проведение эндоскопа до конца толстой кишки. Основным анатомический ориентир завершения процедуры – купол слепой кишки. Полноту исследования по международным стандартам оценивают показателем CIR (Cecal Intubation Rate). Однако визуальное выделение этой области затруднено из-за схожести купола слепой кишки с другими структурами кишечника (складки, бугры и т.п.) [1-2], особенно специалистами с малым опытом.

Для повышения процента полноты колоноскопических исследований разрабатываются системы автоматического обнаружения купола слепой кишки на видеопоследовательностях. Создание таких систем требует комплексного подхода, объединяющего алгоритмы глубокого обучения для детектирования области интереса, методы анализа движения и постобработку данных. В последнее время методы глубокого обучения нашли применение как в медицине, так и в эндоскопических исследованиях. В частности, совмещение нейросетевых детекторов [3-4] с пространственно-временным анализом [5], а также классификацией движения и ситуационной логикой может повысить точность результата.

Таким образом, актуальной задачей является разработка и тестирование комплексного алгоритма обнаружения купола слепой кишки (КСК) на колоноскопических видеоизображениях. В данном исследовании комплексный алгоритм включает использование метода Хорна–Шунка для расчета оптического потока [5], предобученную глубокую нейронную сеть [6-7], алгоритм классификации движения и блок постобработки областей-кандидатов на основе алгоритма ситуационной логики.

Рассмотренная задача является малоизученной для использования систем искусственного интеллекта в эндоскопии, в связи с чем, существует лишь

ограниченное число научных публикаций по обнаружению КСК [3, 8]. Отсутствие унифицированного размеченного датасета, к сожалению, не позволяет провести сравнение результатов на данном этапе разработки.

## II. ОПИСАНИЕ НАБОРА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ

Набор видеопоследовательностей для тестирования алгоритма предоставлен эндоскопическим отделением Ярославской областной клинической онкологической больницы. Всего было собрано 127 видеопоследовательностей исследований толстой кишки. В рамках дальнейшей оценки работы алгоритма была проведена экспертная разметка области видео, содержащей купол слепой кишки [2]. Основные характеристики набора:

- средняя длительность – 17 мин 28 с;
- частота кадров – 25 кадр/секунду;
- разрешение после предобработки – 640×640 пикселей;
- эндоскопическая система – FUJIFILM ELUXEO 7000, OLYMPUS EXERA III (серии 180, 190), PENTAX EPK-i7010.

## III. КОМПЛЕКСНЫЙ АЛГОРИТМ АНАЛИЗА

Алгоритм расчёта оптического потока, предложенный Хорном и Шунком, основывается на гипотезах постоянства яркости и гладкости поля вектора скоростей [5]. Пусть:

$$I(x, y, t): \Omega \times T \rightarrow R, (x, y) \in \Omega \subset R^2, t \in T$$

где  $I(x, y, t)$  – яркость пикселя в точке  $(x, y)$  в момент времени  $t$ .

Для двух последовательных кадров  $I_0(x, y) = I(x, y, t)$ ,  $I_1(x, y) = I(x, y, t+1)$  вектор оптического потока в точке  $(x, y)$  обозначим:

$$w(x, y) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix},$$

где  $u$  и  $v$  – горизонтальная и вертикальная компоненты смещения.

Предположение постоянства яркости:

$$I(x+u, y+v, t+1) \approx I(x, y, t).$$

С разложением Тейлора первого порядка приходим к уравнению оптического потока:

$$I_x u + I_y v + I_t = 0,$$

где  $I_x, I_y, I_t$  – частные производные яркости.

Так как уравнение содержит две неизвестные функции  $u$  и  $v$ , вводится дополнительное условие гладкости, что приводит к задаче минимизации функционала энергии для восстановления поля  $w$ :

$$E(u, v) = \iint_{\Omega} \left[ (I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right] dx dy,$$

где,  $\alpha > 0$  – коэффициент регуляризации.

Минимизация эквивалентна решению системы уравнений Эйлера-Лагранжа:

$$\nabla I(\nabla I w + I_t) - \alpha^2 \Delta w = 0,$$

где  $\nabla I = (I_x, I_y)^T$ ,  $\Delta$  – оператор Лапласа.

Для приближенного решения используется итерационная схема:

$$w^{(k+1)} = \bar{w}^{(k)} - \frac{\nabla I(\nabla I \bar{w}^{(k)} + I_t)}{\alpha^2 + I_x^2 + I_y^2},$$

где  $\bar{w}^{(k)}$  – локальное усреднение  $w^{(k)}$ .

Пространственные и временные производные приближаются сверткой с дискретными масками:

$$k_x = \frac{1}{4} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, k_y = \frac{1}{4} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, k_t = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} (I_1 - I_0).$$

Для регуляризации используются оператор сглаживания с маской:

$$k_1 = \begin{bmatrix} 1/12 & 1/6 & 1/12 \\ 1/6 & 0 & 1/6 \\ 1/12 & 1/6 & 1/12 \end{bmatrix}.$$

Для визуализации векторное поле  $w(x, y)$  представляется в полярных координатах:

$$\rho = \sqrt{u^2 + v^2}, \varphi = \arctan 2(v, u).$$

Таким образом, для отображения значений пикселей используется цветовая модель HSV:

$$H = \frac{\varphi}{2\pi} \cdot 255, S = 255, V = \min(\rho, 1) \cdot 255.$$

Определим алгоритм определения направления движения оптического потока. Пусть входными данными являются тензоры компонент оптического потока:

$$F_x, F_y \in \mathbb{R}^{H \times W},$$

где  $H$  и  $W$  – высота и ширина кадра. Изображение делится на четыре сектора относительно центральных координат:

$$h_{mid} = \left\lceil \frac{H}{2} \right\rceil, w_{mid} = \left\lceil \frac{W}{2} \right\rceil,$$

Каждый секторов изображения определяется как:

$$S_{i,j} = \left\{ (F_x(x, y), F_y(x, y)) \mid x \in H_j, y \in \mathcal{G}_i \right\}, i, j \in \{0, 1\},$$

Для каждого вектора  $(F_x, F_y)$  вычисляем амплитуду и угол направления:

$$M(x, y) = \sqrt{F_x^2 + F_y^2}, \theta(x, y) = \frac{180}{\pi} \arctan 2(F_x, F_y),$$

Для каждого сектора  $S_{ij}$  определяется средние значения:

$$\overline{M}_{ij} = \frac{1}{|S_{ij}|} \sum_{(x,y) \in S_{ij}} M(x, y), \overline{\theta}_{ij} = \frac{1}{|S_{ij}|} \sum_{(x,y) \in S_{ij}} \theta(x, y),$$

Секторы классифицируются по следующим правилам:

- если  $\overline{M}_{ij} < 0,1$ , то сектор считается статичным  $class(S_{ij}) = stop$ ;
- если  $-45^\circ \leq \overline{\theta}_{ij} < 45^\circ$  или  $45^\circ \leq \overline{\theta}_{ij} \leq 135^\circ$ , то движение классифицируется как «вовне»  $class(S_{ij}) = out$ ;
- если  $135^\circ \leq \overline{\theta}_{ij} \leq 180^\circ$  или  $-180^\circ \leq \overline{\theta}_{ij} \leq -135^\circ$ , а также если  $-135^\circ \leq \overline{\theta}_{ij} \leq -45^\circ$  то направление считается «внутри»  $class(S_{ij}) = in$ .

Каждому сектору присваивается вес равный  $w = \frac{1}{4}$ ,

который суммируется по трем категориям, а затем нормируется на суммарный вес:

$$P_k = \frac{1}{4} \sum_{ij} \mathbf{1}\{class(S_{ij}) = k\}, k \in \{in, out, stop\}.$$

Итоговое правило классификации движения:

$$\text{направление} = \begin{cases} stop, & P_{stop} \geq 0.5, \\ in, & P_{in} > P_{out}, \\ out, & \text{иначе.} \end{cases}$$

На рис. 1 изображена блок-схема итогового алгоритма для обнаружения КСК на видеопоследовательности колоноскопического исследования.

Алгоритм Хорна-Шунка применяется для расчёта оптического потока между последовательными кадрами видеоряда. На основе анализа движения выделяются три состояния: «in», «out» и «stop». Если движение классифицируется как «stop», на исходном кадре выполняется поиск КСК с использованием нейросетевого детектора; в противном случае обработка переходит к следующей паре кадров. Далее найденные области-кандидаты используются в алгоритме ситуационной логики. Для этого детекции объединяются в кластеры, при этом временной промежуток между событиями не превышает 1 секунды. Для каждого кластера вычисляются средняя точность детектирования  $\bar{S}$  и продолжительность  $D$ . Алгоритм ситуационной логики оценивает собой оценку области-кандидаты по следующему критерию:

$$\alpha = w * \bar{S} + (1 - w) \frac{D}{D_{max}},$$

где  $w \in [0, 1]$ , – весовой коэффициент, определяющий баланс между средней точностью и длительностью,  $D_{max}$  – максимальная длительность кластера.

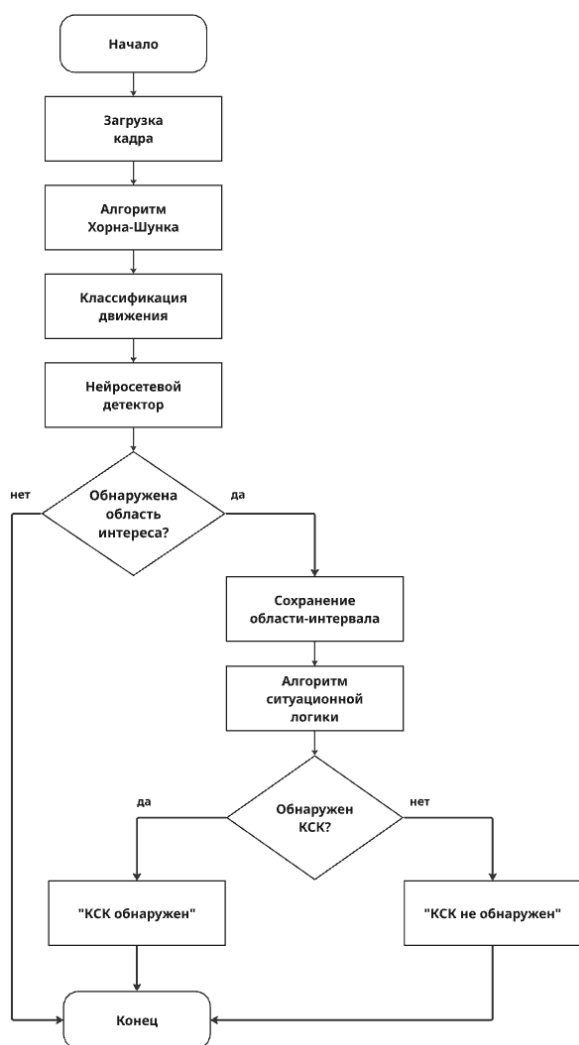


Рис. 1. Блок-схемы работы алгоритмов обнаружения КСК на видеопоследовательности

Если выполняется условие  $\alpha > \alpha_t$ , считается, что КСК присутствует на видеопоследовательности, в противном случае область КСК отсутствует. В рамках проведенного исследования были эмпирически определены параметры  $\alpha_t = 0.4$ ,  $w = 0.6$ .

#### IV. ЧИСЛЕННЫЕ РЕЗУЛЬТАТЫ

Для количественной проверки точности работы комплексного алгоритма анализа эндоскопических видеопоследовательностей используется метрика Intersection over Union (IoU):

$$IoU = \frac{\Delta_{\text{overlap}}}{\Delta_{\text{union}}},$$

где  $\Delta_{\text{overlap}} = \max(0, \min(b, d) - \max(a, c))$  – длительность пересечения интервалов,  $\Delta_{\text{union}} = (b - a) + (d - c) - \Delta_{\text{overlap}}$  – длительность объединения интервалов. Здесь  $a$ ,  $b$  – границы временного отрезка, полученного алгоритмом,  $c$ ,  $d$  – границы, определённые экспертом. Предполагается, что  $a \leq b$  и  $c \leq d$ .

В рамках эксперимента были использованы шесть видеопоследовательностей, каждая из которых предварительно размечена врачом-эндоскопистом с указанием области КСК. Итоговые результаты тестирования алгоритма приведены в табл. 1.

ТАБЛИЦА 1. РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ КОМПЛЕКСНОГО АЛГОРИТМА НА ТЕСТОВЫХ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

Название видеофайла	IoU
Пациент 1	0,25
Пациент 2	0,38
Пациент 3	0,11
Пациент 4	0,21
Пациент 5	0,19
Пациент 6	0,12

Результаты тестирования из табл. 1 для видеофайла Пациент\_2 показали, что комплексный алгоритм достиг наивысшего значения метрики IoU равного 0,38. При этом имеются видеофайлы с меньшими, но отличными от нуля значениями IoU. Низкий показатель метрики в данном случае не свидетельствует о неудовлетворительной работе алгоритма, т.к. временной диапазон предсказанной области КСК значительно меньше, чем диапазон экспертной разметки, что снижает IoU. С практической точки зрения важно, чтобы пересечение областей присутствовало ( $IoU > 0$ ) в пределах окна, определённого экспертом. Подобные случаи нередки, особенно если предсказанная область охватывает лишь несколько кадров, что делает итоговый временной диапазон, вычисленный алгоритмом, значительно меньше экспертного.

#### V. ЗАКЛЮЧЕНИЕ

Результаты исследования подтверждают целесообразность применения комплексного алгоритма для анализа колоноскопических видеоданных. Тестирование показало его устойчивость в определении купола слепой кишки с максимальной метрикой IoU равной 0,38. Подход демонстрирует перспективность использования как системы поддержки врачебных решений. Дальнейшие исследования будут направлены на оптимизацию работы комплексного алгоритма в режиме работы реального времени, а также расширение тестовой базы для дальнейших расширенных испытаний.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Zhou S.K., Greenspan H., Shen D. Deep Learning for Medical Image Analysis. Elsevier Science, 2017.
- [2] Хрящев В.В., Завьялов Д.В., Андержанова А.С. Классификация эндоскопических изображений устья червеобразного отростка на основе методов глубокого машинного обучения // Цифровая обработка сигналов. 2023. № 1. С. 35-38.
- [3] Cho M., et al. Identification of cecum time-location in a colonoscopy video by deep learning analysis of colonoscope movement // PeerJ. 2019. Vol. 7. P. e7256.
- [4] Liu J., Subramanian K.R., Yoo T.S. An optical flow approach to tracking colonoscopy video // Computerized Medical Imaging and Graphics. 2013. Vol. 37. № 3. P. 207-223.
- [5] Horn B.K.P., Schunck B.G. Determining optical flow // Artificial Intelligence. 1981. Vol. 17. P. 185-203.
- [6] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C.-Y., Berg A. SSD: Single Shot MultiBox Detector // European Conference on Computer Vision. Springer, 2016. P. 21-37.
- [7] Sohan M., Sai Ram T., Rami Reddy C.V. A review on YOLOv8 and its advancements // International Conference on Data Intelligence and Cognitive Informatics. Singapore: Springer, 2024. P. 529-545.
- [8] M. Taghiakbari, S. H. Ghalehjegh, E. Jehanno Automated Detection of Anatomical Landmarks During Colonoscopy Using a Deep Learning Model // Journal of the Canadian Association of Gastroenterology, 2023.