IV. Кибербезопасность. Особенности обеспечения безопасности в социальных сетях

Шустов С.А., Мещеряков Р.В.

Повышение безопасности видеоданных за счет использования помехоустойчивого метода видеостеганографии на основе глубоких нейронных сетей

Аннотация: В работе предложен метод скрытой передачи данных в видео, обеспечивающий высокое визуальное качество и надежное извлечение сообщения после типовых искажений. Архитектура основывается на совместном обучении энкодера и декодера; для повышения устойчивости применяется дифференцируемый блок искажений, имитирующий шум, размытие, изменение яркости и масштабирование. Подход позволяет встраивать до 256 бит информации в кадр при среднем PSNR около 31 дБ и BER < 1%. Эксперименты подтверждают скрытность и робастность метода относительно современных аналогов.

Ключевые слова: видеостеганография, скрытая передача данных, устойчивость к искажениям, глубокое обучение, нейросетевой автоэнкодер, PSNR, SSIM, BER

Введение

Стеганография в видеоданных решает задачу скрытой передачи информацию. Например, для канала передачи информации «камера–экран» характерны межкадровые зависимости и неизбежные искажения (масштабирование, шум, муар), что усложняет задачу и требует баланса между незаметностью, устойчивостью и вместимостью [1-3]. Традиционные методы (LSB, модификации DCT/DWT) обеспечивают приемлемое качество только при малой нагрузке и резко теряют надежность при реальных искажениях.

Нейросетевые подходы сокращают этот разрыв: машинное обучение позволяет автоматически подбирать устойчивые способы

кодирования. Для изображений показали эффективность автоэнкодерные архитектуры и дифференцируемые имитаторы шумов (HiDDeN) [2]. Для видео предложены решения с вниманием и GAN-сопровождением (RivaGAN) [3] и мультиуровневые схемы (DVMark) [4], нацеленные на повышение устойчивости при сохранении визуального качества.

Обзор связанных работ

HiDDeN демонстрирует базовый принцип: энкодер, noise-layer и декодер обучаются совместно, оптимизируя одновременно скрытность и точность извлечения [2]. RivaGAN использует механизм внимания и состязательное обучение для повышения устойчивости к сжатию видео [3]. DVMark распределяет полезную нагрузку по пространственно-временным масштабам, повышая переносимость к совокупности искажений [4].

Архитектура модели

Система состоит из энкодера E и декодера D. На вход подаются кадр-контейнер I и сообщение M фиксированной длины L. Энкодер формирует стего-кадр $I' = E(I, M)$, визуально близкий к I. Декодер по искаженному стего-кадру I'' восстанавливает $M = D(I'')$. В энкодере последовательно используются 2D-свертки (извлечение пространственных признаков) и тонкая 3D-свертка 3×1×1 для согласования соседних кадров (скользящее окно из трех кадров), что уменьшает межкадровое мерцание.

Сообщение кодируется полносвязным блоком в карту признаков $k \times H \times W$ и конкатенируется с признаками изображения. Финальная 2D-свертка восстанавливает стего-кадр исходного размера $H \times W \times C$. Декодер – несколько Conv2D-блоков и полносвязный слой на L выходов, которые интерпретируются как вероятности битов (после пороговой обработки).

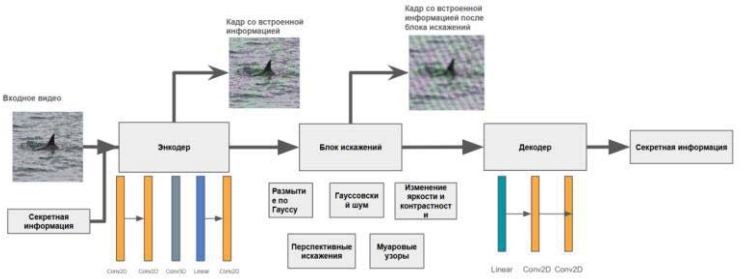

Рисунок 1 – Архитектура предлагаемой модели

Дифференцируемый блок искажений

Модуль $T(\cdot)$, расположенный между E и D, имитирует искажения стего-кадра на этапе обучения. Случайная композиция операций добавляется к каждому мини-батчу: аддитивный гауссов шум, размытие, изменение яркости и контрастности, масштабирование с последующей интерполяцией до исходного размера, а также небольшие аффинно-перспективные преобразования. Операции выбраны дифференцируемыми, что позволяет E обучаться устойчивому внедрению.

Функция потерь:

$$L = \lambda \cdot L_{emb} + L_{msg} + \mu \cdot L_{temp} + \nu \cdot L_{mask},$$

где L_{emb} – MSE между I и I' (скрытность);

L_{msg} – бинарная кроссэнтропия на битах;

L_{temp} – штраф за межкадровое мерцание;

L_{mask} – перцептивное взвешивание ошибок;

λ, μ, ν подбирались эмпирически.

Рисунок 2 – Пример работы дифференцируемого блока искажений

Экспериментальная методика

В качестве данных использовались видеоклипы из набора UCF-101; обучение проводилось в среде PyTorch с оптимизатором Adam ($\alpha = 10^{-4}$) в течение 50 эпох при размере батча 16. На каждом шаге формировалось новое бинарное сообщение длиной L = 256 бит, которое внедрялось в кадры разрешения 256×256; одновременно к полученным стего-кадрам применялась случайная композиция аугментаций из заранее заданного набора. Качество скрытого встраивания оценивалось метриками PSNR и SSIM между $\boldsymbol{I}$ и $\boldsymbol{I}'$ (незаметность), тогда как надежность канала характеризовалась битовой ошибкой BER по всем позициям сообщения и долей экземпляров, восстановленных без каких-либо ошибок.

Сравнение с аналогами

По вместимости предлагаемый метод (256 бит/кадр) существенно превосходит HiDDeN и RivaGAN, а также опережает DVMark (≥100 бит/кадр), сохраняя высокую устойчивость. По качеству изображения ($PSNR \approx 31$ дБ) он уступает RivaGAN и DVMark ($\approx 35 - 37$ дБ), что отражает осознанный компромисс в пользу большей нагрузки при приемлемой незаметности. По устойчивости разработанный подход сопоставим с DVMark и превосходит HiDDeN; при этом у RivaGAN «высокая» устойчивость в первую очередь относится к воздействию сжатия.

Таблица 1 – Сравнение методов видеостеганографии

Метод	Вместимость	PSNR, дБ	Устойчивость
HiDDeN	30–40 бит/кадр	≈36	Средняя
RivaGAN	~30 бит/кадр	≈35	Высокая (к сжатию)
DVMark	≥100 бит/кадр	≈37	Очень высокая
Предлагаемый метод	256 бит/кадр	≈31	Очень высокая

Результаты

Без искажений декодер извлекает сообщение без ошибок (BER = 0 %), при этом средние по видео значения составляют $PSNR \approx 32$ дБ и $SSIM \approx 0{,}92$.

При воздействии типовых искажений – аддитивный гауссовский шум ($\sigma = 0{,}1$), небольшие повороты ($1° - 5°$) и эмуляция съемки экрана (муар) – метод сохраняет робастность: средний BER не превышает $1\,\%$, а при более мягких условиях ошибок не фиксируется; суммарно метрики остаются в допустимых пределах $PSNR \geq 30$ дБ, $SSIM \geq 0{,}85$, $BER \leq 1\,\%$.

На качественных примерах различия между I и I′ визуально минимальны; после искажений заметны артефакты кадра, но декодер извлекает сообщение корректно. Баланс между скрытностью и устойчивостью подтвержден на разнообразных сценах.

Заключение

Предложен нейросетевой метод видеостеганографии, обеспечивающий скрытность и надежное извлечение сообщения после типовых искажений. На «чистых» кадрах достигается BER = 0 % при PSNR ≈ 32 дБ и SSIM ≈ 0,92; при умеренных помехах сохраняются пороговые значения PSNR ≥ 30 дБ, SSIM ≥ 0,85, BER ≤ 1 % (нагрузка 256 бит/кадр). Метод применим для защиты контента и скрытой связи в видеопотоках.

Литература:

1. *Baluja S.* Hiding Images in Plain Sight: Deep Steganography // Advances in Neural Information Processing Systems. – 2017. – P. 2069–2079.

2. *Zhu J., Kaplan R., Johnson J., Fei-Fei L.* HiDDeN: Hiding Data with Deep Networks // ECCV. – 2018. – LNCS 11215. – P. 682-697.

3. *Zhang K.A., Xu L., Cuesta-Infante A., Veeramachaneni K.* Robust Invisible Video Watermarking with Attention (RivaGAN) // arXiv:1909.01285. – 2019.

4. *Luo X., Li Y., Chang H., Liu C., Milanfar P., Yang F.* DVMark: A Deep Multiscale Framework for Video Watermarking // IEEE Transactions on Image Processing. – 2023. – Volume 34. – P. 4371-4385. – DOI: 10.1109/TIP.2023.3251737.

5. *Nah S., Baik S., Hong S., et al.* NTIRE 2019 Challenge on Video Deblurring and Super-Resolution: Dataset and Study / 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). – IEEE: Long Beach, CA, USA, 2019. – DOI: 10.1109/CVPRW.2019.00251.

6. *Soomro K., Zamir A.R., Shah M.* UCF101: A Dataset of 101 Human Action Classes from Videos in the Wild // arXiv:1212.0402. – 2012.

7. *Tancik M., Mildenhall B.; Ng R.* StegaStamp: Invisible Hyperlinks in Physical Photographs. – URL: https://ieeexplore.ieee.org/document/9156548 (дата обращения 1.09.2025).

8. *Hayes J., Danezis G.* Generating Steganographic Images via Adversarial Training // NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. – NY: Curran Associates Inc., 2017. – P. 1951-1960.

9. *Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P.* Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE Transactions on Image Processing. – 2004. – Volume 13, Issue 4. – P. 600-612. – DOI: 10.1109/TIP.2003.819861.

10. *Goodfellow I., Pouget-Abadie J., Mirza M., et al.* Generative Adversarial Nets. – URL: https://archive.org/details/arxiv-1406.2661 (дата обращения 10.09.2025).

Михалевич И.Ф., Пчелинцев Д.И.

Использование регулярных выражений для управления информационной безопасностью интеллектуальных транспортных систем

Аннотация: Статья посвящена проблеме неполноты информации об инцидентах информационной безопасности интеллектуальных транспортных систем. Рассмотрены методы обработки и анализа данных из открытых источников, содержащих фрагментарную информацию о возможных угрозах функционирования