Известия Алтайского государственного университета. 2026. № 1 (147). С. 137–142.
Izvestiya of Altai State University. 2026. No 1 (147). P. 137–142.

Научная статья
УДК: 519.688
https://doi.org/10.14258/izvasu(2026)1-20

Вычислительные методы цифровой стеганографии

Сергей Александрович Шустов[1], Роман Валерьевич Мещеряков[2]

[1]Московский физико-технический институт (национальный исследовательский университет), Москва, Россия, shustov.sa@phystech.edu
[2]Институт проблем управления им В.А. Трапезникова РАН, Москва, Россия, meshcheryakov.roman@gmail.com

Original article

Computational Methods of Digital Steganography

Sergei A. Shustov[1], Roman V. Meshcherya

[1]Moscow Institute of Physics and Technology (National Research University), Moscow, Russia, shustov.sa@phystech.edu
[2]V.A. Trapeznikov Institute of Control Sciences RAS, Moscow, Russia, meshcheryakov.roman@gmail.com

Аннотация. Представлен обзор современных методов цифровой стеганографии для фото- и видеоконтента. Настоящая работа нацелена на разработку нового метода скрытой передачи данных в видеопоследовательностях с учетом временной согласованности. Предложен нейросетевой алгоритм скрытой передачи данных в видеопотоках. Архитектура вычислительного метода объединяет 2D- и 3D-свертки: первые сохраняют детали кадра, вторые усредняют признаки трех соседних кадров, устраняя межкадровое «мерцание». Данные встраиваются покадрово с использованием скользящего окна, поэтому в памяти хранится лишь триада кадров, что облегчает потоковую обработку. При обучении модели используется дифференцируемый блок искажений, обеспечивающий устойчивость к реальным помехам при пересъемке. Составная функция потерь сочетает маскированный MSE (незаметность встраивания), бинарную кросс-энтропию (точность декодирования) и штраф за временную несогласованность. При нагрузке 256 бит на кадр метод достигает на UCF-101 (128×128) PSNR≈31 дБ, SSIM≈0,92 и BER<1 % даже при сильных синтетических искажениях.

Ключевые слова: стеганография, алгоритм, цифровой водяной знак, функция потерь, свертка

Abstract. The paper reviews current digital steganography techniques for photo and video content. It introduces a neural-network algorithm for covert data transmission in video streams. The architecture of computational methods blends 2D and 3D convolutions: the 2D layers preserve fine frame details, while the 3D layer averages features across three neighboring frames to suppress inter-frame «flicker». Data are embedded frame-by-frame via a sliding window, so only a trio of frames needs to reside in memory, enabling efficient streaming. During training, a differentiable distortion module is used to make the system robust against real-world rerecording noise. The composite loss combines masked MSE (embedding imperceptibility), binary cross-entropy (decoding accuracy), and a temporal inconsistency penalty. The proposed method with a payload of 256 bits per frame attains PSNR≈31 dB, SSIM≈0.92.

Keywords: steganography, algorithm, digital watermark, loss function, convolution

Для цитирования: Шустов С.А., Мещеряков Р.В. Вычислительные методы цифровой стеганографии // Известия Алтайского государственного университета. 2026. № 1 (147). С. 137–140. https://doi.org/10.14258/izvasu(2026)1-20.

For citation: Shustov S.A., Meshcheryakov R.V. Computational Methods of Digital Steganography. *Izvestiya of Altai State University*. 2026. No 1 (147). P. 137–140. (In Russ.). https://doi.org/10.14258/izvasu(2026)1-20.

Введение

Под стеганографией понимают науку о скрытой передаче информации [1, с. 5]. Современная стеганография отличается от классической тем, что секретными являются не алгоритмы, а параметры их настройки. Основные задачи стеганографии включают защиту авторских прав, контроль целостности данных и создание скрытых каналов связи. Цифровая стеганография занимается внедрением скрытых данных в цифровые изображения, аудио и видео. Также развивается сетевая стеганография, скрывающая данные в сетевых протоколах (например, в задержках ICMP-пакетов [2, с. 44; 3, с. 470]).

За последнее десятилетие разработано множество методов стеганографии. Одним из первых нейросетевых подходов стало использование сверточной нейросети для скрытия целого изображения внутри другого [4, с. 2070]; позднее появились архитектуры для встраивания битовых сообщений, такие как HiDDeN [5, с. 683], достигающие высокого качества и помехоустойчивости, а совсем недавно был предложен метод RoSteALS [6, с. 933] для устойчивого сокрытия данных. Однако стеганография в видео пока малоизучена.

Настоящая работа нацелена на разработку нового вычислительного метода скрытой передачи данных в видеопоследовательностях с учетом временной согласованности, обеспечивающего минимально заметное искажение видео и высокую точность извлечения сообщения даже после различных искажений.

1. Свойства визуальных данных, влияющие на незаметность встраивания

Для незаметного внедрения информации необходимо учитывать особенности человеческого зрения. Согласно закону Вебера, минимально заметное увеличение яркости пропорционально исходному уровню: на ярком фоне требуется большее приращение, чтобы изменение стало заметным [1, с. 36]. Кроме того, глаз адаптируется к уровню освещенности, и после определенной точки насыщения дальнейшее увеличение яркости не различается (эффект адаптации яркости [1, с. 43]). Также человек менее чувствителен к цветовым изменениям, чем к яркостным, поэтому скрытые данные часто встраивают в цветовые каналы. Важна и контрастная чувствительность: способность заметить изменение зависит от контраста сигнала и фона. Сложный (текстурированный или шумный) фон маскирует встроенный сигнал — мелкие изменения на нем менее заметны. Зрительная система менее восприимчива к искажениям в высокочастотных областях (краях, мелких деталях) и более чувствительна на гладких участках, поэтому искажения следует вносить там, где они наименее заметны.

2. Критерии качества и незаметности стеганографического встраивания

Для количественной оценки незаметности скрытого встраивания используются метрики качества изображения. PSNR (peak signal-to-noise ratio), измеряемый в децибелах, выше 30 дБ обычно означает высокое визуальное качество стего-изображения. Другая метрика — SSIM (structure similarity) — отражает структурное сходство между оригиналом и измененным изображением и более коррелирует с восприятием человека. В стеганографии важен также показатель ошибки декодирования, например BER (bit error rate) — доля неправильно извлеченных бит. Оптимальный алгоритм должен обеспечивать высокие PSNR/SSIM при минимальном BER.

Незаметность встраивания оценивается тем, насколько стего-контейнер визуально неотличим от исходного. Искажения нужно распределять пропорционально локальной чувствительности: минимальные — на однородных участках, более сильные — в областях с деталями. Для этого применяется маска высоких частот (например, фильтр Собеля [7, с. 4]), которая снижает «штраф» за изменение пикселей на границах и текстурах по сравнению с однородными областями.

3. Постановка задачи

Пусть дан входной кадр цветного видео I (например, 128×128 RGB) и двоичное сообщение M длиной k бит, которое нужно скрыто встроить в этот кадр. Требуется обучить стеганографический автоэнкодер, состоящий из энкодера, декодера и вспомогательного блока искажений, чтобы при преобразовании $I \to I' \to I'' \to \hat{M}$ выполнялись условия:

1) минимальное искажение кадра-носителя при получении стего-кадра I';

2) надежное восстановление сообщения $\hat{M} \approx M$ из искаженного стего-кадра I'';

3) согласованность изменений между соседними кадрами видео.

Для достижения этих целей вводится составная функция потерь и специальная архитектура нейросети, учитывающая временную составляющую.

4. Архитектура модели

Разработанная модель [8, с. 2] представляет собой стеганографический автоэнкодер с тремя основными компонентами (см. рис. 1).

1. Энкодер — нейросеть, внедряющая сообщение в кадр. Сначала два последовательных слоя Conv2D извлекают признаки из входного изображения. Затем для учета времени три последовательных кадра объединяются трехмерным слоем Conv3D, который агрегирует информацию соседних кадров и сглаживает

межкадровые флуктуации. Параллельно двоичный вектор сообщения M пропускается через полносвязный слой и преобразуется в карту признаков, совместимую с признаками изображения. Эта карта конкатенируется с признаками кадра, после чего заключительный слой Conv2D восстанавливает число каналов и выдает стего-кадр I', содержащий визуально незаметное скрытое сообщение.

2. Дифференцируемый блок искажений — имитирует возможные помехи при передаче видео (например, пересъемку на камеру). В процессе обучения между энкодером и декодером вставляется слой Distortion, добавляющий к I' случайные искажения: размытие (Gaussian blur), слабый аддитивный шум, случайное изменение яркости/контраста и небольшие геометрические трансформации (масштабирование, поворот). Все эти преобразования дифференцируемы, что позволяет учитывать их влияние при обучении. Реальное видео может подвергаться схожим или даже более сильным искажениям, поэтому задача сети — научиться внедрять сообщение так, чтобы декодер смог его извлечь даже после таких помех.

3. Декодер — нейросеть, которая извлекает сообщение из искаженного кадра I''. Декодер представляет собой каскад сверточных слоев Conv2D и выходной полносвязный слой из нейронов. Каждый выходной нейрон оценивает вероятность того, что соответствующий бит равен 1. После функции активации эти вероятности бинаризуются порогом 0.5, формируя восстановленную битовую строку $\hat{M}$. Декодер обучается совместно с энкодером, стремясь к тому, чтобы $\hat{M} = M$.

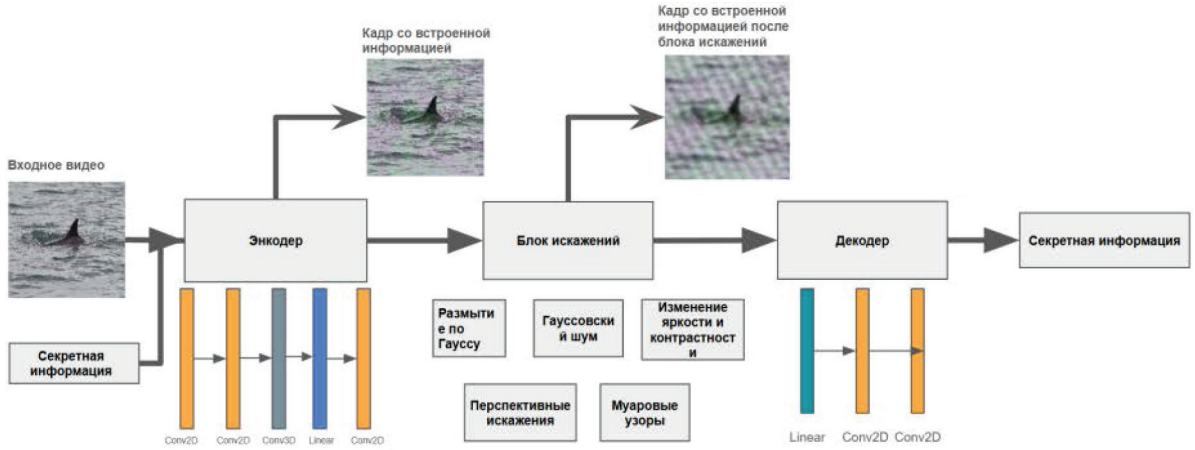

Рис. 1. Схема стеганографического автоэнкодера: энкодер получает исходный кадр I и сообщение M, выдает стего-кадр I'; слой Distortion вносит искажения $I' \rightarrow I''$; декодер восстанавливает сообщение $\hat{M}$

5. Функция потерь

Обучение модели происходит посредством минимизации составной функции потерь L_{total}, которая складывается из трех компонент.

1. Потери встраивания L_{enc} — среднеквадратичная ошибка (MSE) между исходным и стего-кадром (I, I'). Чтобы сильнее штрафовать искажения на однородных участках изображения (где они более заметны), MSE рассчитывается с маской высоких частот (фильтр Собеля [7, с. 4]), поэтому изменения на гладких поверхностях дают больший вклад в ошибку, чем на границах и деталях.

2. Потери декодирования L_{dec} — бинарная кросс-энтропия между исходным сообщением M и предсказанным декодером $\hat{M}$. Минимизация этой компоненты повышает точность декодирования скрытых данных.

3. Временной штраф L_{temp} — ошибка согласованности между соседними кадрами, рассчитанная как MSE между разностями карт встраивания смежных кадров. Он штрафует эффект «мерцания» — резкие скачки яркости или цвета от кадра к кадру. Минимизируя этот штраф, модель обеспечивает изменения плавными во времени.

Общая функция потерь представляет собой взвешенную сумму:

$$L_{total} = \lambda_1 L_{enc} + \lambda_2 L_{dec} + \lambda_3 L_{temp},$$

где веса λ_i подбираются экспериментально для балансировки важности критериев. В вычислительных экспериментах выбрано соотношение $\lambda_1 : \lambda_2 : \lambda_3 = 0.8 : 1 : 0.4$, при котором достигается компромисс между незаметностью и надежностью передачи.

6. Обработка видео (скользящее окно)

Разработанная модель применяется к видеопоследовательности покадрово по принципу скользящего окна. На каждом шаге берется тройка последовательных кадров I_{t-1}, I_t, I_{t+1}; энкодер встраивает сообщение

в центральный кадр I_t с учетом соседних кадров, выдавая стего-кадр I_t'. Затем окно сдвигается: рассматривается следующая тройка I_t, I_{t+1}, I_{t+2}, где I_{t+1} теперь будет центральным кадром для встраивания и так далее. Такой подход позволяет обрабатывать поток практически любой длины без загрузки всего видео в память, сохраняя преимущество учета временной согласованности благодаря 3D-свертке на каждом шаге. На стороне приемника декодер аналогично извлекает сообщение из каждого кадра, зная, что при кодировании учитывались его соседи (для декодирования достаточно одного стего-кадра I_t', так как вся информация для восстановления M содержится в нем, а контекст использовался лишь при встраивании).

7. Экспериментальная настройка

В работе был использован датасет UCF-101. Все видео приведены к разрешению 128×128 (24 кадра/с); из них сформированы тройки кадров для скользящего окна. В каждый кадр скрывается последовательность длиной k = 256 бит (например, случайный набор или фрагмент текста), т.е. нагрузка 256 бит на кадр. Для обучения использован оптимизатор Adam [9] с начальным шагом 10^{-4}, 80 эпох, батч из 8 троек кадров. Для ускорения сходимости и предотвращения переобучения применялись стандартные приемы (снижение шага на плато, малый L2-регуляризатор и др.). Обучение выполнено на GPU; использование дифференцируемых искажений и расчета сложной функции потерь увеличило время одной эпохи, но повысило устойчивость модели к помехам.

8. Результаты

Полученные стего-кадры демонстрируют высокое качество (PSNR ≈ 31 дБ, SSIM ≈ 0.92, минимально заметные искажения). Вероятность битовой ошибки BER менее 1 % даже после всех искусственных помех блока Distortion. Стего-видео визуально практически неотличимо от исходного (рис. 2, табл.). Декодер надежно извлекает скрытое сообщение; ошибки возникают лишь в отдельных битах при самых сильных искажениях.

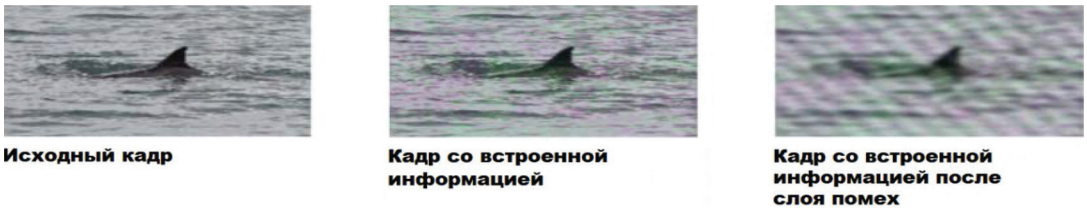

Рис. 2. Пример работы алгоритма встраивания. Слева: исходное изображение, в центре: стего-изображение без искажений, справа: стего-изображение после применения блока искажений

Сравнение разработанного метода с существующими стеганографическими подходами (емкость указана в битах на изображение)

Метод	Вместимость	PSNR (дБ)	Устойчивость	Применимость к видео
HiDDeN [5]	30–40 бит	≈36	Средняя	Нет
StegaStamp [10]	50 бит	≈30	Очень высокая	Нет
RoSteALS [6]	>100 бит	≈32	Высокая	Нет
Разработанный	256 бит	≈31	Очень высокая	Да

Как видно из таблицы, предложенный метод обладает существенно большей вместимостью скрытого канала (256 бит/кадр), остается конкурентоспособным по качеству (PSNR около 31 дБ) и единственный применим к видео. Например, HiDDeN [5, с. 683] дает лучшее качество (около 36 дБ), но встраивает лишь до 40 бит и рассчитан только на статичные изображения. StegaStamp [10, с. 2114] чрезвычайно помехоустойчив (выдерживает даже печать и фотографирование), однако ограничен емкостью 50 бит и также предназначен для единичных кадров. RoSteALS [6] передает более 100 бит при PSNR≈32 дБ, обеспечивая высокую устойчивость за счет сокрытия данных в латентном пространстве автоэнкодера, но не предназначен для видео.

Заключение

В работе рассмотрены основы цифровой стеганографии и предложен метод скрытой передачи данных в видеопоследовательностях с использованием глубоких нейронных сетей. Модель позволяет встраивать до 256 бит в каждый кадр цветного видео, обеспечивая качество стего-видео PSNR≈31 дБ, SSIM≈0.9 и надежное извлечение данных при отсутствии значительных помех. Использование узкого временного контекста и дифференцируемых иска-

жений обеспечивает устойчивую передачу данных при минимально заметных артефактах. Метод закладывает основу для практической системы видеостеганографии, пригодной для широкого спектра приложений (защита медиаконтента, аутентификация данных и др.). В дальнейшем планируется повысить устойчивость алгоритма к более сильным реальным искажениям (например, сжатие кодеком), оптимизировать скорость работы модели и реализовать адаптивные сценарии с переменной нагрузкой.

Библиографический список

1. Федосеев В.А., Митекин В.А. Теоретические основы стеганографии и цифровых водяных знаков. Самара: Изд-во Самарского ун-та, 2017. 130 с.

2. Johnson N.F., Katzenbeisser S. A Survey of Steganographic Techniques // Information Hiding Techniques for Steganography and Digital Watermarking / red. S.A. Katzenbeisser, F. P. Petitcolas. Boston: Artech House, 2000. P. 43–78.

3. Chan C.K., Cheng L.M. Hiding Data in Images by Simple LSB Substitution // Pattern Recognition. 2004. Vol. 37. № 3. P. 469–474. https://doi.org/10.1016/j.patcog.2003.08.007

4. Baluja S. Hiding Images in Plain Sight: Deep Steganography // Advances in Neural Information Processing Systems 30 (NeurIPS 2017). Red Hook, NY: Curran Associates, Inc., 2017. P. 2069–2079.

5. Zhu J., Kaplan R., Johnson J., Li F.-F. HiDDeN: Hiding Data with Deep Networks // Computer Vision — ECCV 2018. Lecture Notes in Computer Science, vol. 11129. Cham: Springer, 2018. P. 682–697. https://doi.org/10.1007/978-3-030-01267-0_40

6. Bui T., Agarwal S., Yu N., Collomosse J. RoSteALS: Robust Steganography Using Autoencoder Latent Space // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2023). Los Alamitos, CA: IEEE Computer Society, 2023. P. 933–942. https://doi.org/10.1109/CVPRW59228.2023.00100

7. Sobel I., Feldman G. A 3×3 Isotropic Gradient Operator for Image Processing. Stanford: Stanford Artificial Intelligence Laboratory, 1968. 10 p. ResearchGate. URL: https://www.researchgate.net/publication/239398674_A_3x3_Isotropic_Gradient_Operator_for_Image_Processing (дата обращения: 09.08.2025).

8. Шустов С.А., Мещеряков Р.В. Модель стеганографического встраивания информации в изображения методами глубокого обучения // Информационные процессы. 2025. Т. 25. № 1. С. 1–13. https://doi.org/10.53921/18195822_2025_25_1_1

9. Kingma D.P., Ba J.L. Adam: A Method for Stochastic Optimization // International Conference on Learning Representations (ICLR 2015), San Diego, CA. Ithaca, NY: Cornell University, 2015. ArXiv e-print. URL: https://arxiv.org/abs/1412.6980 (дата обращения: 09.08.2025).

10. Tancik M., Mildenhall B., Ng R. StegaStamp: Invisible Hyperlinks in Physical Photographs // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020). Los Alamitos, CA: IEEE Computer Society, 2020. P. 2114–2123. https://doi.org/10.1109/CVPR42600.2020.00219

References

1. Fedoseev V.A., Mitekin V.A. *Theoretical Foundations of Steganography and Digital Watermarks*. Samara: Izdatelstvo Samarskogo Universiteta, 2017. 130 p. (In Russ.).

2. Johnson N.F., Katzenbeisser S. A. Survey of Steganographic Techniques. *Information Hiding: Techniques for Steganography and Digital Watermarking*. Ed. by S.A. Katzenbeisser, F.A.P. Petitcolas. Boston: Artech House, 2000. P. 43–78.

3. Chan C.K., Cheng L.M. Hiding Data in Images by Simple LSB Substitution. *Pattern Recognition*. 2004. Vol. 37. No. 3. P. 469–474. https://doi.org/10.1016/j.patcog.2003.08.007

4. Baluja S. Hiding Images in Plain Sight: Deep Steganography. *Advances in Neural Information Processing Systems (NeurIPS 2017)*. Red Hook, NY: Curran Associates, Inc., 2017. P. 2069–2079.

5. Zhu J., Kaplan R., Johnson J., Li F.-F. HiDDeN: Hiding Data with Deep Networks. *Computer Vision (ECCV 2018). Lecture Notes in Computer* Science, Vol. 11129. Cham: Springer, 2018. P. 682–697. https://doi.org/10.1007/978-3-030-01267-0_40

6. Bui T., Agarwal S., Yu N., Collomosse J. RoSteALS: Robust Steganography Using Autoencoder Latent Space. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2023)*. Los Alamitos, CA: IEEE Computer Society, 2023. P. 933–942. https://doi.org/10.1109/CVPRW59228.2023.00100

7. Sobel I., Feldman G. *A 3×3 Isotropic Gradient Operator for Image Processing.* Stanford: Stanford Artificial Intelligence Laboratory, 1968. 10 p. ResearchGate. URL: https://www.researchgate.net/publication/239398674_A_3x3_Isotropic_Gradient_Operator_for_Image_Processing (accessed: 09.08.2025).

8. Shustov S.A., Meshcheryakov R.V. A Robust Steganographic Model for Data Embedding into Images Using Deep Neural Networks. *Information Processes.* 2025. Vol. 25. No. 1. P. 1–13. (In Russ.). https://doi.org/10.53921/18195822_2025_25_1_1

9. Kingma D.P., Ba J.L. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations (ICLR 2015)*, San Diego, CA. Ithaca, NY: Cornell Univer-

sity, 2015. ArXiv e-print. URL: https://arxiv.org/abs/1412.6980 (accessed: 09.08.2025).

10. Tancik M., Mildenhall B., Ng R. StegaStamp: Invisible Hyperlinks in Physical Photographs. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020)*. Los Alamitos, CA: IEEE Computer Society, 2020. P. 2114–2123. https://doi.org/10.1109/CVPR42600.2020.00219

Информация об авторах

С.А. Шустов, магистрант кафедры инфокоммуникационных систем и сетей, Московский физико-технический институт (Национальный исследовательский университет), Москва, Россия;

Р.В. Мещеряков, доктор технических наук, профессор, главный научный сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия.

Information about the authors

S.A. Shustov, Master Student of Department of Infocommunication Systems and Networks, Moscow Institute of Physics and Technology (National Research University), Moscow, Russia;

R.V. Meshcheryakov, Doctor of Sciences in Technology, Professor, Chief Researcher, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia.