

На правах рукописи

Чечуров Александр Александрович

**Исследование методов применения нейронных сетей для обработки  
видео и изображений**

Направление подготовки  
09.04.01 Информатика и вычислительная техника  
направленность – Инженерия программного обеспечения и информационных  
систем  
программа академической магистратуры

**АВТОРЕФЕРАТ**  
магистерской диссертации  
на соискание квалификации (степени) магистра

Екатеринбург 2025

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет телекоммуникаций и информатики» Уральский технический институт связи и информатики (филиал) в г. Екатеринбурге (УрТИСИ СибГУТИ)

Научный руководитель



**Новиков Сергей Николаевич**

Доктор технических наук, профессор кафедры  
ИСТ УрТИСИ СибГУТИ

Рецензент

**Кусайкин Дмитрий Вячеславович**

Кандидат технических наук, доцент кафедры  
МЭС УрТИСИ СибГУТИ,

Защита состоится «30» июня 2025 г. в 9 часов в Федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет телекоммуникаций и информатики» Уральский технический институт связи и информатики (филиал) в г. Екатеринбурге (УрТИСИ СибГУТИ), г. Екатеринбург, ул. Репина, д. 15.

Секретарь Государственной аттестационной комиссии

## Общая характеристика работы

### Актуальность исследования

Современные задачи восстановления, преобразования и анализа визуального контента становятся всё более важными в таких сферах, как цифровая реставрация, медицина, системы видеонаблюдения и мультимедиа. Особое внимание уделяется задачам колоризации – восстановлению цветовой информации на основе чёрно-белых изображений или видеок кадров. До недавнего времени они решались вручную или с использованием простых алгоритмических правил. Однако такие методы оказались неэффективны при работе с большими объёмами данных и нестандартизированными изображениями.

Развитие нейросетевых архитектур, в частности моделей глубокого обучения, дало возможность решать задачи колоризации с высокой степенью автоматизации и реалистичности. В этой связи особенно актуально изучение архитектур, способных восстанавливать цветовую информацию с сохранением контекста сцены и временной согласованности, особенно в задачах обработки видео.

### Объект исследования

Процессы автоматической колоризации и восстановления изображений и видеопотоков с помощью нейросетевых моделей.

### Предмет исследования

Методы и архитектуры глубокого обучения, применяемые для восстановления цвета изображений и видеоданных.

### Цель работы

Разработка и реализация нейросетевой модели, способной автоматически восстанавливать цветовую информацию в изображениях и видеопотоках на основе чёрно-белого входа.

### Задачи исследования

1. Провести обзор существующих подходов и архитектур нейронных сетей для обработки изображений и видео.
2. Исследовать методы предобработки визуальных данных и формирования обучающих пар.
3. Выбрать и адаптировать архитектуру для задачи колоризации.
4. Реализовать программный модуль колоризации на языке Python с использованием PyTorch.
5. Обучить модель на датасете пар ч/б и цветных изображений.
6. Разработать пайплайн обработки видео с покадровой цветовой реконструкцией.
7. Оценить устойчивость модели к искажениям и провести тестирование.

### Научная новизна

1. Адаптация архитектуры U-Net под задачу колоризации изображений и видео, с ориентацией на реализм и устойчивость к искажениям.

2. Реализация единого программного пайплайна, включающего предобработку, инференс и обратную сборку видео.
3. Проведение сравнительного анализа устойчивости модели к типичным искажениям: шуму, сжатию, падению яркости и др.
4. Разработка и тестирование метода визуального и количественного контроля качества результатов.

Теоретическая значимость данной работы обусловлена вкладом в развитие методов применения нейросетевых архитектур в области восстановления и генерации визуальных данных. В исследовании глубоко проанализированы существующие подходы к архитектурам глубокого обучения, в частности – U-Net, GAN, трансформерные и диффузионные модели. Работа расширяет представление о возможностях их адаптации к задаче колоризации, отличающейся высокой неоднозначностью и требующей реконструкции утраченной информации на основе косвенных признаков.

Особое внимание уделяется выделению особенностей колоризации как задачи контекстно-зависимой регрессии, отличной от классификации и сегментации. Автор подробно рассматривает, какие параметры моделей критичны для точного восстановления цветовой информации, как влияет структура признаков на стабильность предсказаний, и каким образом можно повысить обобщающую способность модели путём архитектурных и обучающих модификаций.

Дополнительный вклад в теорию вносит рассмотрение устойчивости моделей к искажениям входных данных и формализация критериев, по которым оценивается робастность нейросетевых систем. В исследовании предлагаются методы построения таких систем с высокой переносимостью, что актуально для задач компьютерного зрения в нестабильных условиях.

Таким образом, теоретическая значимость заключается в систематизации знаний о применении архитектуры U-Net для генеративных задач, адаптации её к визуальной реконструкции, анализе метрик и разработке стратегии обучения, направленной на устойчивость и воспроизводимость результата.

Практическая значимость работы определяется тем, что в ходе выполнения была создана и протестирована реальная, функционально завершённая система, способная автоматически восстанавливать цветовую информацию по чёрно-белым изображениям и видеокадрам. Разработанная модель может быть легко интегрирована в существующие рабочие процессы в различных отраслях, что делает её высоко востребованной в прикладной практике.

Во-первых, система применима для цифровой реставрации архивных фото- и видеоматериалов, включая кинохронику, документальные записи, музейные фонды и библиотеки. Ручная колоризация в таких случаях является крайне ресурсоёмкой, в то время как автоматизированное решение существенно ускоряет процесс, позволяя восстанавливать утраченные визуальные образы в историческом контексте.

Во-вторых, модель может использоваться в системах видеонаблюдения и безопасности, где необходимо восстановить цветовую информацию с камер, работающих в монохромном режиме (например, ИК-камеры в ночных условиях).

В-третьих, в сфере медицинской визуализации, где цвет может нести дополнительную диагностическую ценность, система способна быть частью предварительной или вспомогательной обработки изображений. Особенно это актуально в случаях, когда оригинальные данные имеют ограниченную палитру или не передают цвет в полной мере.

Также система может быть встроена в графические и видеоредакторы, использована в образовательных продуктах, на онлайн-платформах, и как модуль для расширения функционала уже существующих нейросетевых решений.

Высокая модульность реализации, ориентация на открытые фреймворки (PyTorch, OpenCV) и гибкость архитектуры позволяют масштабировать систему под различный уровень задач и объём данных – от настольного приложения до облачного сервиса.

Методологической основой работы является системный подход к решению задачи восстановления визуальных данных, сочетающий методы компьютерного зрения, теории искусственного интеллекта и архитектурного проектирования нейросетевых моделей. Исследование базируется на совмещении теоретического анализа (литературный обзор, классификация моделей, обоснование выбора архитектуры) с экспериментальной реализацией полного цикла решения прикладной задачи.

В рамках работы применены следующие методы:

- архитектурное проектирование нейросетей на базе u-net с модификацией выходных слоёв;
- препроцессинг данных: нормализация, преобразование цветовых пространств, масштабирование и формирование обучающих пар;
- обучение модели с использованием стохастического градиентного спуска, оптимизатора adam, а также функций потерь mse и перцептивных критериев (feature loss);
- регуляризация и аугментация: dropout, случайные трансформации, увеличение устойчивости модели к шуму и артефактам;
- оценка качества: применение метрик ssim, mse, визуального сравнения, тестирование на устойчивость к искажениям;
- реализация пайплайна обработки видеоданных: покадровая обработка, сбор выходного видеофайла, синхронизация кадров.

В ходе выполнения проекта использовались инструменты: Python, PyTorch, OpenCV, NumPy, Matplotlib, а также open-source датасеты. Вся методика направлена на достижение высокой точности, воспроизводимости результатов и адаптируемости системы к различным условиям.

Достоверность результатов подтверждается многими факторами, как на этапе теоретического анализа, так и в экспериментальной части работы.

Прежде всего, все используемые методы опираются на современные научные подходы и широко признанные в профессиональном сообществе инструменты и технологии. Основные архитектурные решения базируются на открытых публикациях и проверенных научных разработках.

Реализация модели и последующие тестирования были выполнены на реальных данных, включая изображения и видеоконтент, не входившие в обучающую выборку. Это позволило проверить устойчивость модели к ранее не встречавшимся сценариям.

Кроме того, в работе проведён многоступенчатый анализ качества результатов – как по объективным численным метрикам (SSIM, MSE), так и по субъективному визуальному восприятию. Были протестированы условия с типовыми искажениями (шум, сжатие, размытость), и зафиксирована стабильность модели в различных условиях.

Также была реализована проверка промежуточных результатов по эпохам, визуализация предсказаний, логирование ошибок и параметров, что усилило прозрачность обучения и позволило отслеживать динамику прогресса модели.

Достоверность результатов обеспечивается:

- теоретическим обоснованием архитектурных и обучающих решений;
- воспроизводимостью экспериментов;
- устойчивыми показателями на тестовых выборках;
- чёткой логикой построения и анализа нейросетевой модели.

Объём и структура работы

Работа включает введение, три главы, заключение, список литературы из 40 источников и приложения. Общий объём – 85 страниц, включая 1 рисунок, 41 листингов и 12 таблиц промежуточных результатов.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

### Введение

Во введении обоснована актуальность применения нейросетевых архитектур в задачах автоматической обработки изображений и видео, в том числе для реконструкции цветовой информации. Сформулированы цель, задачи, объект и предмет исследования. Приведена гипотеза, обоснована научная новизна и практическая значимость проекта. Также изложены методологические подходы: анализ архитектур глубокого обучения, выбор модели, методы обучения, предобработки данных и оценки качества результатов. Объём введения позволяет получить общее представление о направлениях и значении работы.

### **В первой главе «Анализ научных и технических подходов к обработке изображений и видео с использованием нейросетей»**

рассмотрены основные подходы к автоматической обработке визуального контента, включая архитектуры свёрточных нейронных сетей, автоэнкодеры, GAN, трансформеры и диффузионные модели. Проанализированы публикации 2018–2025 гг., касающиеся архитектур для восстановления изображений, цветовой реконструкции и генерации визуального контента. Выделены ключевые особенности различных архитектур, их плюсы и минусы в контексте задач колоризации. Также представлен обзор задач компьютерного зрения, связанных с восстановлением цвета: сегментация, суперразрешение, inpainting и колоризация. Проведена систематизация известных методов предобработки данных, аугментации, нормализации, и описаны метрики оценки качества (SSIM, MSE, PSNR и др.). Сделан вывод о том, что существующие решения требуют адаптации под задачу автоматической цветовой реконструкции, особенно в видеопотоке, где важно сохранять временную согласованность.

### **Во второй главе «Теоретические основы и архитектурные решения нейросетевой колоризации»**

рассматриваются ключевые аспекты проектирования модели: выбор архитектуры, структура входных и выходных данных, преобразование цветových пространств, подходы к обучению и регуляризации. Особое внимание уделено архитектуре U-Net, как одной из наиболее подходящих для задач генеративной реконструкции. Приведено обоснование выбора U-Net по сравнению с другими подходами (GAN, трансформеры, автоэнкодеры), описана её структура и механизм skip-соединений.

Детально рассмотрены этапы предобработки изображений и видео: ресайз, нормализация, выделение яркостных каналов, преобразование RGB → Lab, формирование обучающих пар.

Также описан метод обучения модели, включая функции потерь, оптимизаторы, регуляризацию и аугментацию. Подробно рассмотрены механизмы устойчивости к шуму, искажениям и нестандартным входам. Приведены сценарии расширения модели на видео: покадровая обработка,

использование ConvLSTM и 3D-свёрток. Глава формирует прочную теоретическую базу для практической реализации.

### **В третьей главе «Разработка, реализация и тестирование модели цветовой реконструкции изображений и видео»**

представлены этапы проектирования и реализации программного модуля. Подробно описана реализация архитектуры U-Net на языке Python с использованием библиотеки PyTorch. Раскрыты аспекты конфигурации модели: количество уровней, фильтры, активации, батч-нормализация. Показана реализация пайплайна: от загрузки и предобработки изображений и кадров видео до вывода предсказаний и сохранения результатов. Описаны примеры кода и их объяснения.

Отдельно выделен этап тестирования: мониторинг ошибок, логирование, визуализация промежуточных результатов, оценка метрик. Выполнена проверка устойчивости модели к искажениям (шум, сжатие, падение яркости, размытие), показано, как это влияет на результат.

Описан модуль обработки видеопотока: извлечение кадров, покадровая обработка, реконструкция финального видео. Представлены итоги оценки качества и визуальные выводы, подтверждающие эффективность модели. Результаты продемонстрировали устойчивую работу на новых данных.

## **ЗАКЛЮЧЕНИЕ**

В заключении подведены итоги исследования, подтверждена реализация поставленной цели и решение всех задач. Сформулированы ключевые выводы по теоретической и практической части, дана оценка применимости полученного решения. Приведены перспективы развития — расширение архитектуры, внедрение внимания и временных связей, переход к онлайн-реализации. Работа показала успешность применения нейросетевых методов для восстановления цветовой информации в изображениях и видео, а также высокую адаптивность модели к сложным условиям.

**Основные теоретические и практические результаты работы заключаются в комплексной реализации подхода к задаче автоматической цветовой реконструкции визуальных данных с использованием нейросетевых методов.** В теоретической части работы был выполнен обширный обзор существующих архитектур глубокого обучения, в частности, свёрточных нейросетей, автоэнкодеров, генеративно-состязательных моделей, трансформеров и диффузионных подходов. На основании анализа было обосновано использование архитектуры U-Net, обладающей способностью восстанавливать пространственные признаки и эффективно передавать информацию от входного уровня к выходному. В процессе разработки была адаптирована модель U-Net к задаче колоризации изображений и видео, модифицирована под восстановление цветowych каналов в пространстве Lab и реализована с использованием фреймворка PyTorch.

На практике был создан полноценный программный модуль, включающий все этапы обработки: от подготовки и предобработки данных до

вывода результатов и обратной сборки видеопотока. В ходе обучения модели применялись современные методы регуляризации и аугментации, а также функции потерь, направленные на повышение точности и реалистичности цвета. Результаты тестирования подтвердили работоспособность системы как при обработке статичных изображений, так и при обработке видео. Дополнительно была проведена серия экспериментов по оценке устойчивости модели к различным типам искажений, включая шум, сжатие, снижение освещённости и размытость, что позволило зафиксировать стабильность и адаптивность полученного решения к условиям реального применения.

**Практическая значимость исследования определяется высоким прикладным потенциалом разработанного модуля, направленного на автоматическую реконструкцию цвета в визуальных данных.** Полученное решение может быть эффективно использовано в различных отраслях, где требуется обработка изображений и видео: в цифровой реставрации исторических материалов и кинохроники, в повышении качества архивных видеозаписей, в образовательных и мультимедийных продуктах, а также в системах видеонаблюдения и медицинской визуализации. Универсальность и масштабируемость модели позволяют адаптировать её под разнообразные задачи и типы входных данных, а реализация на открытых инструментах делает её пригодной для внедрения как в исследовательских, так и в коммерческих решениях.

Система может выступать как отдельное программное средство, так и как часть более крупного комплекса по обработке визуальной информации. Её устойчивость к внешним искажениям, способность обрабатывать данные в пакетном и покадровом режиме, а также возможность интеграции в веб-интерфейсы и облачные сервисы подтверждают актуальность и высокую применимость результатов, полученных в рамках данной магистерской диссертации.

## **СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ**

А. А. Чечуров, Л. Н. Евдакова

### **Применение нейронных сетей для обработки видео в реальном времени: проблемы и решения**

Министерство цифрового развития, связи и массовых  
коммуникаций РФ Уральский технический институт связи и  
информатики (филиал) ФГБОУ ВО  
«Сибирский государственный университет телекоммуникаций и информатики» в  
г. Екатеринбурге (УрТИСИ СибГУТИ), Россия

В статье рассмотрены проблемы и решения, связанные с применением нейронных сетей для обработки видео в реальном времени. Описаны ключевые аспекты, такие как проблемы задержки, вычислительных ресурсов и качества обработки, которые ограничивают возможности использования нейронных сетей в реальных условиях.

Приведены подходы для оптимизации этих процессов, включая методы параллельных вычислений, ускорения работы нейронных сетей с помощью специализированного оборудования (GPU, TPU), также рассмотрены примеры успешных решений в области видеонаблюдения, мониторинга и анализа видео, которые обеспечивают высокую точность и минимальную задержку в обработке данных. Проведен анализ существующих технологий и их применения в различных отраслях, таких как безопасность, медицина и развлечения, что позволяет оценить текущий потенциал и перспективы использования нейронных сетей для обработки видео в реальном времени.

Ключевые слова: нейронные сети, обработка видео, реальное время, задержка, параллельные вычисления, GPU, TPU, видеонаблюдение, мониторинг, анализ видео, оптимизация.

## **Application of neural networks for real-time video processing: problems and solutions**

Ministry of Digital Development, Communications and Mass  
Communications of the Russian Federation Ural Technical Institute of Communications  
and  
Informatics (branch)  
"Siberian State University of Telecommunications and Informatics" in  
Yekaterinburg (UrTISI SibGUTI), Russia

The paper discusses the challenges and solutions associated with the application of neural networks for real-time video processing. Key aspects such as latency, computational resources, and processing quality issues that limit the use of neural networks in real-world applications are described. Approaches to optimize these processes are presented, including parallel computing techniques, acceleration of neural networks using specialized hardware (GPU, TPU), and examples of successful solutions in video surveillance, video monitoring and analysis that provide high accuracy and minimal latency in data processing. The existing technologies and their applications in various industries such as security, medicine and entertainment are analyzed, which allows us to assess the current potential and prospects of using neural networks for real-time video processing.

Keywords: neural networks, video processing, real time, latency, parallel computing, GPU, TPU, video surveillance, monitoring, video analysis, optimization.

### **Введение**

В последние годы нейронные сети значительно трансформировали различные области, включая обработку видео в реальном времени. Их способность извлекать сложные паттерны и особенности из визуальных данных открывает новые возможности для множества приложений, таких как видеонаблюдение, анализ поведения, автомобильные системы помощи водителю и развлекательные технологии. Использование нейронных сетей в реальном времени позволяет не только улучшить точность распознавания объектов, но и обеспечить быструю реакцию на изменения в видеопотоке, что имеет критическое значение для многих высокотехнологичных систем [1].

Применение нейронных сетей для обработки видео в реальном времени сталкивается с рядом проблем, таких как высокие вычислительные требования, задержки в обработке и необходимость адаптации моделей к ограниченным ресурсам. Эти проблемы требуют разработки эффективных методов оптимизации, таких как использование специализированного аппаратного обеспечения (GPU, TPU) и создание легких нейронных

архитектур. Важным аспектом является также балансировка между точностью результатов и скоростью обработки, что критично для реальных приложений. В данной статье будут рассмотрены основные вызовы, с которыми сталкиваются современные системы обработки видео в реальном времени, а также предложены возможные решения для их преодоления.

### **Главные проблемы для применения нейронных сетей для обработки видео в реальном времени**

Одной из основных проблем применения нейронных сетей для обработки видео в реальном времени является высокая вычислительная нагрузка, связанная с обработкой больших объемов данных. Видеопотоки содержат огромное количество информации, которая требует значительных вычислительных мощностей для извлечения признаков и классификации объектов. Для эффективной работы нейронных сетей в таких условиях часто используются графические процессоры (GPU) или специализированные ускорители (TPU), что помогает повысить скорость обработки. Однако даже с использованием таких устройств обработка видео может быть слишком медленной для ряда приложений, например, в автономных транспортных системах или видеонаблюдении, где важна минимальная задержка. Таким образом, задача минимизации времени обработки без потери качества остаётся одной из самых актуальных.

Также есть необходимость в оптимизации нейронных сетей для работы в реальном времени при сохранении точности. Обычно более сложные модели нейронных сетей обеспечивают более высокую точность распознавания объектов, но требуют гораздо больше вычислительных ресурсов и времени на обработку. Это создает дилемму между точностью и скоростью, которую необходимо решить для успешной интеграции нейронных сетей в реальное время [2].

Не менее важной проблемой является адаптация моделей нейронных сетей к разнообразию входных данных, таких как различные разрешения видео, шумы и изменения условий освещенности. Видео, получаемое в реальных условиях, часто бывает подвержено искажениям, что может значительно ухудшить качество обработки и снизить точность распознавания объектов. Кроме того, модели должны быть устойчивыми к изменениям в окружающей среде, что требует использования методов повышения общей стабильности и адаптивности сетей. Все эти факторы требуют разработки более универсальных и гибких решений, которые могут эффективно справляться с разнообразием условий в реальном времени.

### **Способы решения главных проблем при применении нейронных сетей**

Одним из способов решения проблем высокой вычислительной нагрузки и задержек при обработке видео в реальном времени является использование специализированного аппаратного обеспечения. Графические процессоры (GPU) и тензорные процессоры (TPU) значительно ускоряют процесс обработки, обеспечивая параллельную обработку данных и ускоряя выполнение сложных вычислений. Также для оптимизации применяются методы распределенной обработки, которые позволяют использовать вычислительные ресурсы нескольких устройств одновременно. Использование данных технологий снижает нагрузку на отдельные компоненты системы, позволяя обрабатывать видеопотоки с минимальными задержками. Можно также применять методы аппаратного ускорения, такие как использование специализированных чипов для обработки нейронных сетей, что делает возможным реализацию высокоскоростных и энергоэффективных систем обработки видео в реальном времени [3].

Другим эффективным методом решения проблемы высокой вычислительной нагрузки является разработка легких и оптимизированных архитектур нейронных сетей. Сетевые модели, такие как MobileNet, SqueezeNet или EfficientNet, предназначены для

работы на устройствах с ограниченными ресурсами, позволяя сохранять высокую точность распознавания при значительном уменьшении вычислительных затрат. Эти архитектуры обычно используют меньшее количество параметров и слоев, что позволяет повысить скорость обработки и снизить потребление энергии. Важно, что такие решения находят широкое применение в мобильных и встроенных системах, где ресурсы ограничены, но при этом требуется быстрая обработка видео с высокой точностью.

### **Методы повышения устойчивости нейронных сетей к шуму в изображениях и видео**

Одной из основных проблем, с которой сталкиваются нейронные сети при обработке изображений и видео, является чувствительность к шуму, который может значительно ухудшить качество результатов. Шум в изображениях и видео может возникать по разным причинам, включая технические ограничения камер, плохие условия освещенности, а также компрессию данных. Это приводит к искажению исходных данных и снижению точности распознавания объектов или событий. Для повышения устойчивости нейронных сетей к шуму разработаны различные методы, которые направлены на улучшение качества входных данных и минимизацию влияния искажений.

Один из методов повышения устойчивости – использование предварительной обработки данных. Это может включать фильтрацию изображений и видео с помощью традиционных методов, таких как медианные или гауссовы фильтры, которые позволяют снижать уровень шума перед подачей данных в нейронную сеть. Также широко применяются методы денойзинга с использованием нейронных сетей, такие как автоэнкодеры, которые обучаются восстанавливать чистые изображения из зашумленных. Данные подходы позволяют значительно улучшить качество входных данных, что в свою очередь способствует повышению точности распознавания и минимизации влияния шума [4].

Другим подходом является разработка более устойчивых архитектур нейронных сетей, которые могут эффективно обрабатывать зашумленные данные. Применение таких архитектур, как сверточные нейронные сети (CNN) с дополнительными слоями для обработки шумов, позволяет сети обучаться распознавать объекты в условиях, когда присутствует значительное количество шума. В некоторых случаях используются специальные регуляризационные методы, такие как Dropout или Batch Normalization, которые помогают нейронным сетям более эффективно адаптироваться к шуму, улучшая обобщающие способности модели. Такие методы делают сеть менее чувствительной к случайным искажениям и способствуют повышению ее устойчивости.

Кроме того, активно исследуются подходы, связанные с обучением нейронных сетей на зашумленных данных. Методики обучения с шумом (Noise Robust Training) включают создание обучающих наборов с добавленным искусственным шумом, что позволяет модели обучаться на более сложных и разнообразных данных. Также применяются методы увеличения данных, которые включают в себя генерацию различных вариантов изображений с добавлением шума для улучшения общего качества обучения. Таким образом, нейронные сети становятся более устойчивыми к шуму, улучшая свои показатели на реальных данных, где уровень шума может существенно варьироваться [5].

### **Методы адаптации моделей нейронных сетей к разнообразию входных данных**

Один из методов адаптации нейронных сетей к разнообразию входных данных заключается в применении техники увеличения данных (data augmentation). Данный подход включает создание множества вариаций исходных изображений или видеок кадров с целью повышения универсальности модели. Увеличение данных может включать изменения таких параметров, как повороты, изменения масштаба, сдвиги, изменение яркости и

контраста, а также добавление шума. Это позволяет нейронной сети обучаться на более разнообразном наборе данных, что способствует повышению её устойчивости к изменениям в условиях, таких как различные освещенности, ракурсы или другие вариации исходных данных. Таким образом, модель становится более универсальной и способной адаптироваться к новым, непредсказуемым данным [6].

Еще одним подходом к адаптации является использование техники трансферного обучения. В этом случае нейронная сеть обучается на большом объеме данных из одной области, а затем её знания переносятся и адаптируются к новой области с меньшим количеством данных. Это особенно полезно, когда для обучения модели в новой области недостаточно данных. Трансферное обучение позволяет использовать предварительно обученные модели и адаптировать их под специфические условия, например, для работы с определёнными типами видео или изображений, где могут присутствовать необычные объекты или фоны. С помощью этого подхода можно эффективно решать задачи, связанные с разнообразием входных данных, не начиная обучение с нуля.

Кроме того, для повышения адаптивности нейронных сетей к различным типам данных применяются методы регуляризации и адаптивного обучения. Например, использование слоев Batch Normalization помогает улучшить обобщающую способность модели, ускоряя её обучение и делая её более устойчивой к различным условиям. Методы, такие как Dropout, также способствуют предотвращению переобучения и улучшению способности модели к обобщению, что важно при наличии разнообразных данных. Адаптивные алгоритмы обучения, такие как Adam или AdaGrad, позволяют более эффективно настраивать параметры модели в зависимости от особенностей входных данных, что улучшает её производительность в условиях неопределенности или изменчивости данных.

## **Заключение**

Применение нейронных сетей для обработки видео в реальном времени представляет собой важную и перспективную область, которая находит широкое применение в таких сферах, как видеонаблюдение, автономные транспортные системы, медицина и развлечения. Однако, несмотря на очевидные преимущества, существует ряд значительных проблем, связанных с вычислительными затратами, задержками в обработке, а также с влиянием шума и разнообразием входных данных. Решение этих проблем требует комплексного подхода, включающего как оптимизацию нейронных сетей, так и использование специализированного аппаратного обеспечения, что позволяет обеспечить необходимую скорость и точность обработки.

Текущие исследования и разработки в области нейронных сетей для обработки видео в реальном времени обещают значительные улучшения в производительности, точности и устойчивости систем. Применение нейронных сетей в реальном времени продолжит развиваться, предоставляя новые возможности для множества отраслей, от обеспечения безопасности до медицинской диагностики и интеллектуальных транспортных систем. Важно, чтобы дальнейшие исследования фокусировались на усовершенствовании существующих методов и создании более адаптивных и универсальных моделей, способных эффективно работать в условиях ограниченных вычислительных ресурсов и разнообразных входных данных.

## **Список литературы**

1. Гудфеллоу, И., Бенгио, Й., и Курвиль, А. Глубокое обучение. MIT Press // 2016
2. Schmidhuber, J. Глубокое обучение в нейронных сетях: An Overview. Springer. // 2015
3. O'Reilly Media. Глубокое обучение для компьютерного зрения. // 2017

4. Müller, A. C., & Guido, S. Введение в машинное обучение с Python: A Guide for Data Scientists. O'Reilly Media. // 2016
5. Bishop, C. M. Распознавание образов и машинное обучение. Springer. // 2006
6. Zhang, K., & Zhang, Z. Глубокое обучение для компьютерного зрения: Краткий обзор. Wiley. // 2008

Чечуров А.А., МИВТ-31

Научный руководитель: к.п.н., доцент кафедры ЭС Новокшенова Р.Г.

## The methods of training neural networks

Neural networks are a powerful tool in machine learning, and they are trained using various methods depending on the nature of the data and the task at hand. Here, we'll explore three key approaches: supervised learning, unsupervised learning, and reinforcement learning.

### 1. Supervised Learning

In supervised learning, the neural network is trained on labeled data, where the correct answer (label) for each input is already known. The goal is to learn how to predict these labels for new, unseen data [1].

Applications: Image recognition, text processing, classification tasks.

### 2. Unsupervised Learning

In unsupervised learning, the network works with unlabeled data and attempts to uncover hidden patterns, structures, or distributions within the data [2].

Applications: Clustering, dimensionality reduction, segmentation.

### 3. Reinforcement Learning

In reinforcement learning, an agent interacts with the environment, taking actions and receiving rewards or penalties. Based on these rewards, the agent learns to choose the optimal actions to achieve a goal [3].

Applications: Games, robotics, autonomous systems.

### Conclusion

These three training methods – supervised learning, unsupervised learning, and reinforcement learning – form the foundation of modern neural networks and are applied in diverse fields, ranging from data analysis to autonomous systems and artificial intelligence.

Method Chosen: Supervised Learning

Reasons for Choosing Supervised Learning:

Labeled Data: For tasks such as classification, object detection, and segmentation, labeled data is often readily available, such as images and videos.

Convolutional Neural Networks (CNNs): CNNs are ideal for image and video analysis, utilizing algorithms like backpropagation for learning.

Efficiency with Large Datasets: Supervised learning methods can efficiently train models on large datasets of images and video materials.

Final Thoughts: Supervised learning remains one of the most powerful and widely used methods for training neural networks, particularly for tasks involving large datasets and well-defined labels, such as image and video processing.

### Список источников:

1. What is Supervised Learning? [Электронный ресурс] Режим доступа: <https://towardsdatascience.com/maths-behind-supervised-learning-for-dummies-the-theory-in-plain-words-part-i-8f9be4d7e33a> (Дата обращения: 25.11.2024)
2. NFC technology devices [Электронный ресурс] Режим доступа: <https://www.analyticsvidhya.com/blog/2020/04/supervised-learning-unsupervised-learning/> (Дата обращения: 25.11.2024)

3. NFC working [Электронный ресурс] Режим доступа:  
[https://spinningup.openai.com/en/latest/spinningup/rl\\_intro.html](https://spinningup.openai.com/en/latest/spinningup/rl_intro.html) (Дата обращения: 25.11.2024)