

**магистрант: Аубакиров Е.Е., к.ф.-м.н., сениор-лектор: Иманбаев К.С.,  
лектор-магистр: Тұрғынбаева А.А.**

*Алматинский технологический университет, г.Алматы, Республика Казахстан*

## **Архитектура модели сверточной нейронной сети**

### **АННОТАЦИЯ**

Сверточная нейронная сеть – это специальная архитектура искусственных нейронных сетей, которая является частью технологии глубокого обучения, направленной на эффективное распознавание изображений. Сверточная нейронная сеть – является основным средством распознавания объектов, лиц на фотографиях. Сверточная нейронная сеть с помощью специальной операции - позволяет одновременно уменьшить количество информации, хранящейся в памяти, благодаря чему она хорошо работает с изображениями высокого разрешения и выделяет опорные признаки изображения, такие как грани, контуры. Цель статьи рассмотреть архитектуру сверточной нейронной сети и изучить ее связь с наукой глубокого обучения.

**Ключевые слова:** сверточная сеть, искусственный нейрон, глубокое обучение, распознавание изображений, сетевая архитектура.

### **АНДАТПА**

Конволюциялық нейрондық желі – жасанды нейрондық желілердің арнайы архитектурасы болып табылады, ол кескінді тиімді тануға бағытталған, терең оқыту технологиясының бөлігі болып табылады. Конволюциялық нейрондық желі – нысандарды, фотосуреттердегі тұлғаларды танудың негізгі құралы. Арнайы операцияны қолдану арқылы конволюциялық нейрондық желі – жадта сақталатын ақпарат мөлшерін бір уақытта азайтуға мүмкіндік береді,

соның арқасында ол жоғары ажыратымдылықтағы суреттермен жақсы жұмыс істейді және жиектер, контурлар сияқты кескіннің тірек белгілерін ерекшелейді. Мақаланың мақсаты конволюциялық нейрондық желі архитектурасын қарастыру және терең оқыту ғылымымен байланысын зерттеу.

**Негізгі сөздер:** конволюциялық желі, жасанды нейрон, терең оқыту, кескінді тану, желі архитектурасы.

Одной из сложных задач в информационной технологии является обработка и распознавание объектов в изображениях. Актуальность данной проблемы связана с тем, что исследования по распознаванию объектов, анализу образов и речи входят в перечень приоритетных направлений в развитии науки и техники. Процесс распознавания символов на изображениях со сложным фоном можно разделить на три этапа: выбор области расположения символов заданных на изображении, отдельные символы деления и распознавание символов.

В настоящее время такие технологии реализуются тремя традиционными методами-структурным, описательным и шаблонным. Эти методы ориентированы на условия их эффективного использования, но все эти методы имеют свои недостатки. Особенно ярко эти недостатки проявились при масштабном использовании программно-технологических систем. Во всех системах распознавания символов на изображениях характеристики точности резко снижаются, а различные искажения входного изображения становятся технологически приемлемыми. При этом технологические условия получения информации о маркировке не позволяют полностью устранить эти искажения. Анализ методов разметки и распознавания показал, что для решения этой задачи на изображаемых объектах необходимо эффективно использовать искусственные нейронные сети, поскольку они плохо чувствительны к искажениям входного сигнала, а также обеспечивают возможность получения классификатора, который лучше моделирует сложную функцию распределения

изображений символов, тем самым повышая скорость и точность решения задачи другими методами.

Чаще всего в задачах распознавания изображений и идентификации изображений используются классические нейронные сетевые архитектуры (многослойный перцептрон, сети с радиально-базисной функцией и др.), но при анализе этих работ и экспериментальных исследований использование классических нейронно-сетевых архитектур в задаче оказывается неэффективным в силу следующих обстоятельств:

- изображения имеют большой размер, соответственно увеличивается размер нейронной сети;
- емкость системы увеличивается в зависимости от количества параметров, продлевая время процесса;
- для повышения эффективности системы целесообразно использовать несколько нейронных сетей, но это увеличивает вычислительную сложность решения задачи и время выполнения[1].

Поэтому для решения задачи выбора области расположения символов были выбраны сверточные нейронные сети, так как они обеспечивают прямую устойчивость к изменению масштаба. Каждый слой сверточной нейронной сети представляет собой набор элементов, состоящих из нейронов. Нейроны одной плоскости имеют одинаковые синаптические коэффициенты, ведущие ко всем локальным частям предыдущего слоя. Каждый нейрон слоя получает вход из определенной области предыдущего слоя (локальное поле приема), то есть входное изображение предыдущего слоя сканируется через небольшое окно и проходит через набор синаптических коэффициентов, а результат отображается на соответствующем нейроне текущего слоя. Таким образом, набор элементов может быть найден в описательных картах и в разделах каждого плоскостного изображения в любом месте предыдущего слоя.

Сверточные нервные сети считаются одним из самых популярных типов нейронных сетей. Он доказал свою эффективность в распознавании сверточного визуального изображения или образа, в системах представления и

при обработке естественного языка. Сверточная нейронная сеть является одной из моделей глубокого обучения нейронной сети. Преимущества данной модели:

- они отлично масштабируются и могут быть использованы для распознавания любых моделей высокого разрешения;
- они используют объемные (трехмерные) нейроны. Нейроны внутри слоя связаны только небольшими областями, называемыми рецептивным полем;
- нейроны соседних слоев связаны с механизмом пространственной локализации. Наложение многих таких слоев обеспечивает нелинейные фильтры, которые больше реагируют на пиксели;
- каждый фильтр расширяется, чтобы охватить все поле зрения. Эти установки дублируются, делятся доступными настройками и формируют карту. Таким образом, все нейроны определенного сверточного слоя начинают реагировать на один объект (в области реального ответа)[2].

Искусственный интеллект значительно развился на пути сокращения разрыва между возможностями людей и машин. Специалисты в этой области, наряду с энтузиастами, работают во многих аспектах в этой области и в конечном итоге стремятся достичь удивительных результатов. Одним из таких актуальных и важных аспектов является система машинного зрения (машинное зрение). Задача этого направления состоит в том, чтобы научить систему машинного обучения тому, что может сделать человечество, легко и просто принять ее и даже использовать различные знания для решения проблем, таких как распознавание изображений и изображений, анализ и классификация этих изображений, преобразование медиа-материала, система рекомендаций, обработка естественного языка и многое другое. Со временем достижения глубокого обучения в области машинного обучения улучшились в основном благодаря конкретному алгоритму, то есть с помощью этих сверточных нейронных сетей.

Сверточная нейронная сеть – это алгоритм глубокого обучения, способный воспринимать изображение как вход, присваивать различные его аспекты или объекты степени значимости (рассматривающие веса и отклонения) и отличать

одно от другого. В простейших методах, когда фильтры создавались вручную, сверточный нейросеть мог самостоятельно считывать эти фильтры (фильтры) и характеристики при достаточном уровне обучения[3].

Архитектура сверточной сети (ConvNet) похожа на модель связи нейронов человеческого мозга, и ее основой является организация его поля зрения. Отдельные нейроны реагируют на раздражители только в ограниченной области визуального поля, называемой рецептивной. Такие поля перекрываются друг с другом, образуя коллекцию, охватывающую весь спектр. В случае простых двоичных изображений метод может демонстрировать среднюю точность прогнозирования классов, но везде точность сводится к минимуму или даже нулю, когда речь идет о сложных изображениях, зависящих от пикселей. ConvNet может успешно фиксировать пространственные и временные зависимости от изображения с помощью соответствующих фильтров. Эта архитектура хорошо обрабатывает набор данных из-за повторного использования весов и уменьшения количества параметров. Другими словами, сеть будет направлена на обучение лучшему пониманию сложности изображения.

Роль ConvNet заключается в том, чтобы привести изображения в удобный для обработки вид, не теряя важных функций для создания хороших прогнозов. Это важно при проектировании архитектуры, которая масштабируется не только с функциями чтения, но и с большим набором данных. Задача сверточной операции заключается в извлечении из вводного изображения признаков высокого уровня, например, таких функций высокого уровня, как края из изображения. ConvNet не ограничивается сверточным слоем. Условно, первое руководство ConvLayer отвечает за фиксацию знаков низкого уровня на свойствах цвета, направления градиента и т.д. Добавляя дополнительные слои, архитектура также адаптируется к особенностям верхнего уровня, обеспечивая полную человеческую понятную сеть изображений в наборе данных. Эта операция может иметь два типа результатов-в одной функция, которая выходит из сети, уменьшается относительно входа в измерение, а в другом размер либо

увеличивается, либо остается неизменным. Подобно сверточному слою, вспомогательный выборочный слой отвечает за уменьшение пространственного размера сверточной характеристики [4].

ConvNet имеет различные архитектуры, которые являются ключом к созданию алгоритмов и обеспечивают и продолжают работу искусственного интеллекта в ближайшем будущем. Это: LeNet, AlexNet, VGGNet, GoogLeNet, ResNet и ZFNet.

LeNet - это структура сверточной нервной системы, предложенная Янном Лекуном и другими. В целом, LeNet относится к lenet-5 и является простой сверточной нервной системой. Нейронные сети LeNet - это модель продвинутой глубокой нейросети, которая может реагировать на часть окружающих клеток в диапазоне искусственных нейронов и демонстрировать хорошие результаты при широкомасштабной обработке изображений.

AlexNet - это одно из названий сверточной нейронной сети, разработанной Алексом Крижевским вместе с Джеффри Хинтоном. Основным результатом было то, что глубина этой модели была важна для ее высокой производительности, которая была более интенсивной с компьютерной точки зрения, но стала возможной благодаря использованию блоков обработки графики во время тренировки.

VGGNet - это усовершенствованная версия AlexNet, которая заменяет большие фильтры (11 и 5 размеров в первом и втором слоях соответственно) несколькими фильтрами 3x3. Сеть VGGNet обучалась с использованием видеокарт.

GoogLeNet - используется не для его точности, а для эффективности размера модели и необходимого количества вычислений. GoogLeNet использует линейный фильтр по всему размеру, принимает текущее количество измерений и смешивает их линейно, делая их меньше, что оказывается очень эффективным.

ResNet - это искусственная нейронная нервная сетка в форме, основанной на известных структурах из пирамидальных клеток в оболочке мозга.

Остаточные нервные сети делают это, пропуская соединения или ярлыки, чтобы пройти через некоторые слои. Типичные модели ResNet реализуются с двух-или трехуровневыми пробелами, включающими нелинейные значения и структурную нормализацию между ними.

Благодаря архитектуре ZFNet уровень погрешности первых пяти категорий снизился. Здесь ключевую роль играет точная настройка гиперпараметров (размер и количество фильтров, размер пакета, скорость чтения и т.д.)[5].

## **Заключение**

Для существенного повышения эффективности рекомендуется использовать сверточные нейронные сети при выборе и распознавании моделей систем в изображениях со сложным фоном, так как они эффективно подходят для изменения многих функций и других искажений. Создается сверточная нейронная сеть, обеспечивающая обнаружение и разделение области расположения символов в сложных изображениях. Для выделения отдельных символов рекомендуется использовать алгоритм, основанный на построении гистограмм средней интенсивности пикселя. Распознавание отдельных символов осуществляется сверточной нейронной сетью, в которой символы функционируют как классификаторы. Проводится оптимизация структуры разработанных нейронных сетей, что позволяет повысить эффективность их работы.

Работа сверточной нейронной сети обычно интерпретируется как переход от определенных особенностей изображения к более абстрактным частицам, а затем к разделению понятий более высокого уровня в абстрактных частицах. Кроме того, сеть самостоятельно настраивается, отфильтровывая несущественные детали и выделяя важные, создавая необходимую иерархию (последовательность признаков) абстрактных признаков. Сверточная нейронная сеть является одним из лучших алгоритмов распознавания и классификации изображений.

## **Список литературы:**

1. «Сверточная нейронная сеть временного веса для распознавания действий». Достижения в области информационно-коммуникационных технологий IFIP-2008. Чам: международное издательство Springer. С. 97-108.
2. Крижевский, Алекс; Суцкевер, Илья; Hinton, Geoffrey E. (2017-05-24). «Классификация ImageNet с глубокими сверточными нейронными сетями» 84-90.
3. Достижения в системе обработки нейронной информации «An Artificial Neural Network-Temporal Bipolar Patters: Application to Phoneme Classification». 1: 31–40.
4. «Исследование ускорителей на основе FPGA для сверточных нейронных сетей», NCAA, 2018.
5. Jain, V. and Seung, S. H. (2008). Natural image denoising with convolutional networks. In NIPS'2008.