**Тема 1.5 Особенности подхода глубокого обучения в задачах   
компьютерного зрения**

**План**

1. Полносвязная сеть
2. Функция активации
3. Функция потерь
4. Оптимизация параметров

**Полносвязная сеть**

Одной из первых нейронных сетей в современном виде является многослойный персептрон Румельхарта – по-другому полносвязная нейронная сеть. Нейроны представляют структуру типа граф, имеющий в качестве узлов сумматор с обучаемыми коэффициентами и нелинейную функцию – функция активации. Такая структура называется персептрон. По существу, персептрон – это нелинейная регрессия. То есть нейронная сеть - это некоторая структура типа нелинейной регрессии.

**Функция активации**

Первым вариантом функции активации была функция сигмоид. Эта функция получается аналитически при решении задачи логистической регрессии. Однако, сегодня сигмоид используется только при решении задач бинарной или много метоточной классификации. Это связано с некоторыми недостатками функции сигмоид. В частности, производная сигмоида почти везде дает 0 значение. Это приводит к проблеме «вымывание» градиента. В большинстве внутренних слоев используется полу-линейная функция активация ReLU. Напомним, что для задачи многоклассовой классификации используется функция активации softmax. Также важно понимать, что функция ReLU приводят недостаткам в ряде случаев. Поэтому иногда используют модификации ReLU. Модификации могут быть как линейными, например ReLU с насыщением или ReLU с утечкой в области значений меньше нуля. Также есть и нелинейные модификации ReLU, например, ELU, GELU или другие. Как правило, эти нелинейные функции аппроксимируют ReLU почти для всех значений. Однако, в критических областях они пытаются разрешить проблемы классической функции. В целом общая рекомендация начинать с использования ReLU, а другие функции использовать лишь если базовый подход приводит к проблемам.

**Функция потерь**

В зависимости от решаемой задачи должны быть выбрана функция потерь. Функция потерь – это функция, по которой проводится оптимизация параметров нейронной сети. Для разных функций потерь могут быть получены разные решения одной и той же задачи. Поэтому выбор функции потерь – это одна из наиболее сложных и важных проблем. Для задачи бинарной классификации используется логистическая функция потерь или бинарная корсс-энтропия. Эта функция получается аналитически при решении задачи логистической регрессии. В случае многоклассовой регрессии используется функция потерь кросс энтропия. Для решения задачи регрессии используются или функция эвклидова расстояния (сумма квадратов разностей) или ее комбинация с модулем разностей. В ряде случаев этих функций недостаточно. Например, в случае несбалансированной выборки. Позже в курсе будут показаны приемы решения таких задач.

**Оптимизация параметров.**

Кроме функции потерь результат решения задачи может завесить от подхода к оптимизации параметров. Базовый выбор оптимизатора – это стохастический метод градиентного спуска. Этот метод реализует метод обратного распространения ошибки для случайно выбранных подвыборок (батчей). Этот метод часто модифицируют, добавляя к нему скользящее среднее (и импульс) или регуляризацию Тихонова. В целом этот метод не гарантирует достижения глобального минимума функции потерь. Однако, это и не всегда нужно. Часто лучше достичь локального минимума близкого к глобальному если результат при этом будут более стабильных. Это связано с тем, что оптимизация должна давать широкую обобщающую способность. Напомним, что обобщающая способность – это разность между ошибкой на тренировочных и любых тестовых данных. Эта разность должна быть минимальной.

**Стохастический метод градиентного спуска с импульсом** – это базовый метод оптимизации, могут быть и другие методы, в том числе адаптивные. В качестве базового решения задачи оптимизации нейронных сетей часто рекомендуют начинать с адаптивного метода Adam. Однако, при достаточно тщательном подборе параметров стохастического спуска можно получить более высокие результаты чем для Adam.

**Резюме**

Учет и знание особенностей подхода глубокого обучения в задачах   
компьютерного зрения позволяет понимать и сравнивать различные подходы решения задач компьютерного зрения. С практической точки зрения указанные в данном разделе особенности будут влиять на результаты решения.