

1. Какой тип обучения Вы бы использовали для решения задачи подсчета кол-ва объектов на изображении? Ответ обоснуйте.

Обучения с учителем, так как точно известно (по крайней мере можно определить) действительное количество объектов (правильные ответы). При таких условиях использование этого типа целесообразно, хотя и потребует от исследователей подготовки достаточной для выполнения задачи обучающей выборки.

2. Что такое расширенная свертка?

Расширение (дилатация) свертки – один из гиперпараметров сверточного слоя СНС. В расширенных свертках между ячейками располагается пустое пространство (пробелы), размеры которого и зависят от дилатации слоя. Ячейки, таким образом, ограничиваются пустыми (нулевыми) элементами друг от друга.

3. Можно ли применять CNN для обработки звуковых сигналов?

Да, потому что сверточные НС используют, когда есть необходимость распознавания образов (чаще это работа с изображениями). Преимущество СНС в данном случае (с изображениями) заключается в гибкости моделей касательно смещений, поворотов необходимых для определения образов. Т. о., сеть, определяющая, например, госномер автомобиля по фотографии, «заметит» его, в какой бы он части фотографии не оказался. Полагаю, что в случае обработки аудио СНС могут обладать таким же преимуществом перед другими архитектурами сетей, находя необходимый звук в любом месте аудиофайла.

4. Почему выбрана скорость обучения 0.001?

Были проведены тесты с разными показателями скорости от 0.1 до 0.00001. Самые лучшие результаты обучения (низкие потери и высокая точность) достигались при `learning_rate = 0.001`. При 0.1 результаты сильно хуже, позже они улучшались до значения скорости в 0.001, после чего снова стали падать. В результате выбрана скорость обучения, равная 0.001.

5. Чем чревато использование ненормированных данных в данной задаче?

Ненормированные входные данные приведут к слишком большим или слишком маленьким активациям нейронов внутренних слоев ($\gg 1$ или $\ll -1$), что в свою очередь скажется на результатах работы всей сети, так как в таком случае активации нейронов выходного слоя будут близки к 1 или -1. Это сведет активационную функцию выходного слоя (в данном случае `softmax`) к простой ступенчатой. Также это может стать причиной сильной чувствительности сети к небольшим изменениям входных данных, что нежелательно, особенно в задачах распознавания рукописных символов.