

# Реферат по статье SpinalNet: Deep Neural Network with Gradual Input

Михаил Лепехин

## 1 Введение

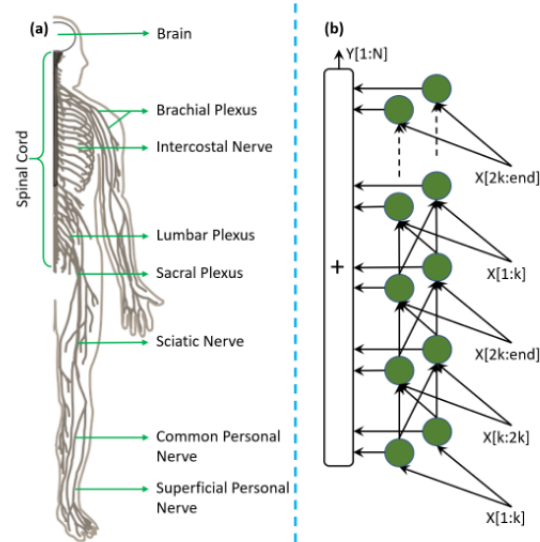
За последние несколько лет глубокие нейронные сети позволили достичь высоких результатов в ряде задач компьютерного зрения. Однако, как правило, это происходило за счёт огромного количества параметров. При этом важно, чтобы размер первого скрытого состояния был достаточно большим. Иначе деградирует обобщающая способность модели. Другая проблема классических глубоких сетей для компьютерного зрения – угасающие градиенты (vanishing gradients).

Авторы SpinalNet предложили способ более экономичного построения нейросетей, вдохновившись строением соматосенсорной системы человека.

Человеческий мозг получает огромное количество информации с помощью тактильных сенсорных нейронов. Хотя точный механизм их работы неизвестен, хорошо доказана ключевая роль нейронов спинного мозга в распознавании ощущений. Некоторое время назад исследователи разработали свёрточные нейронные сети, имитирующие зрительную кору головного мозга кошек, что привело к значительному улучшению качества моделей [1]. Этот успех в моделировании искусственного интеллекта вдохновил авторов на создание сетей с постепенным вводом.

Все сети, использующиеся в компьютерном зрении можно разделить на свёрточные и не-свёрточные. Глубокие свёрточные сети представляют собой комбинации свёрточных и полносвязных слоёв. В большинстве современных архитектур используются пулинги, цель которых снизить число параметров модели. Но у пулингов есть и проблема – они резко сокращают количество информации.

## 2 Описание архитектуры

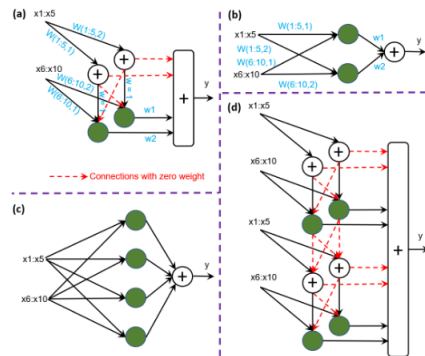


Архитектура разделяет следующие особенности человеческой соматосенсорной системы.

- Постепенный вход и нервное сплетение
- Добровольные и вынужденные действия
- Внимание к интенсивности боли

Структура сети состоит из входного, промежуточного и выходного столбцов.

Главная идея – заменить большие полносвязные слои последовательностью нескольких Spinal слоёв, каждый из которых имеет не очень большой размер. Можно визуальнo показать, почему полносвязные слои могут быть так представлены.



### 3 Эксперименты и результаты

Основная задача, на которой авторы тестировали эффективность и точность SpinalNet, – задача классификации изображений. Они пытались улучшить результаты ряда широко используемых сетей, заменяя полносвязные слои этих сетей на Spinal FC.

На датасете MNIST использование классической CNN со SpinalNet дало точность 98.44%, что превосходит точность просто CNN (98.17%). При использовании SpinalNet с VGG-5 был получен accuracy 99.72, что входит в топ-20 самых высоких результатов полученных когда-либо.

На Fashion-MNIST accuracy при использовании SpinalNet с VGG-5 составила 94.68%. Использование VGG-5 в чистом виде даёт accuracy 94.63%. Таким образом, было получено более высокое качество при помощи более вычислительно эффективной модели.

На датасетах Kuzushiji-MNIST и QMNIST использование SpinalNet с VGG-5 позволило повысить точность исходной VGG-5 с 98.94% до 99.15% и с 99.66% до 99.68% соответственно. Таким образом, на этих датасетах был получен SOTA.

На экспериментах с ResNet точность модели при использовании Spinal FC слоёв упала. Авторы связывают это с затуханием градиента на начальных слоях.

Кроме того, был проведён ряд экспериментов с Transfer Learning. Были взяты модели VGG-19bn и Wide-Resnet-101, предобученные на ImageNet. В дальнейшем на ряде датасетов (CIFAR-10, CIFAR-100, Caltech-101, Bird225, Stanford Cars, SVHN, CINIC-10, STL-10, Oxford 102 Flower, Fruits360) проводился эксперимент с дообучением этих сетей в чистом виде и с заменой полносвязных слоёв на Spinal FC-слои. В итоге на большинстве удалось превзойти результаты исходных сетей, а на Caltech-101, Bird225 и Fruits360 – получить SOTA.

### 4 Заключение

Достаточно убедительно была продемонстрирована эффективность модели SpinalNet. Она требует намного меньше вычислительных ресурсов чем глубокие нейронные сети, при этом её точность лишь незначительно уступает. Кроме того, при комбинации с VGG-5 на 4 датасетах был достигнут state-of-the-art результат. Это доказывает эффективность структуры хордовой нервной системы.

### Список литературы

- [1] D. H. Hubel и T. N. Wiesel. “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”. в: *The Journal of physiology* 160.1 (1962), с. 106–154.