

ИСПОЛЬЗОВАНИЕ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ РАСПОЗНАВАНИЯ ЛИЦ НА ИЗОБРАЖЕНИЯХ

Казиев А. Б.

Научный руководитель: В.Г. Спицын
Томский политехнический университет
abk5@tpu.ru

Введение

Распознавание лиц, детектированных на изображении, на сегодняшний день является актуальной задачей биометрии, решение которой критически важно в широком диапазоне применений. Алгоритмическая обработка изображений с целью выяснения личности детектированного лица применяется в системах общественной безопасности, в системах видеонаблюдения и контроля за перемещениями, в системах контроля и управления доступом к закрытым помещениям и территориям, а также в рамках криминалистической экспертизы.

Одним из активно используемых подходов к этой задаче на сегодняшний день является применение искусственных нейронных сетей особой архитектуры, называемых свёрточными нейронными сетями (СНС), которые позволяют использовать локальные признаки изображений и проявляют более высокую устойчивость к инвариантным преобразованиям распознаваемых объектов по сравнению с другими методами.

В данной работе проводится краткий обзор современных методов применения СНС и результатов, полученных путём применения описанных методов.

Архитектура СНС

Искусственные нейронные сети являются биологически подобными системами, математическая модель которых основана на принципах функционирования головного мозга. СНС являются их дальнейшим развитием, заимствуя принцип использования локальных признаков из нейробиологии. В работе [1] утверждается, что визуальная кора головного мозга состоит из комплекса клеток, каждая из которых чувствительна только к ограниченному участку поля зрения и реагирует на присутствие в собственной области определённых примитивных структур, выполняя роль локальных фильтров для входных данных. СНС эмулирует поведение визуальной коры головного мозга [2].

Архитектура СНС включает в себя свёрточные слои, слои подвыборки и полносвязные слои. Архитектура СНС использует преимущества двумерной структуры изображений с помощью метода локальной связности, ограничивая количество связей между нейронами скрытого свёрточного слоя и входными данными. Конкретно, каждый нейрон скрытого слоя связан только с ограниченным локальным участком изображения.

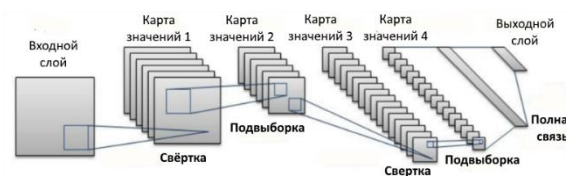


Рис.1. Архитектура свёрточной нейронной сети
СНС использует разделяемые веса так, чтобы каждый нейрон скрытого слоя имел общий набор весов с другими нейронами этого слоя. Веса нейронов скрытого слоя представляют из себя набор фильтров, при помощи которых осуществляется операция свёртки входного изображения. Промежуточными результатами сети являются результаты свёртки каждым отдельным фильтром, называемые картами признаков. Слои подвыборки группируют карты признаков и агрегируют значения, полученные в результате свёртки [3].

Методы распознавания на основе СНС

Значительное количество работ в области распознавания лиц опирается на разработку СНС различной архитектуры и использование их в каскаде с другими СНС или иными распознавателями.

Применение глубоких СНС (состоящих из большого количества слоёв) рассмотрено в работе [2], в которой СНС используется для классификации объектов на базе изображений ImageNet. СНС с глубокой архитектурой также часто используется в задаче распознавания лиц. К примеру, модификация такой архитектуры используется в работе [4], в которой также подробно исследуется процедура подготовки базы лиц для обучения сети. Разработанной авторами глубокой СНС получен результат 98.95% точности на популярной для тестирования СНС базе лиц Labeled Faces in the Wild (LFW) [5].

Увеличение размера СНС как в глубину, так и в ширину подразумевает значительное увеличение количества настраиваемых параметров и затрат вычислительных ресурсов. Стремясь обойти эту проблему, команда исследователей из корпорации Google, совместно с сотрудниками Университета Северной Каролины и Мичиганского университета, создали архитектуру СНС для распознавания объектов под названием Inception [6], в которой использовался принцип нахождения биологически подобной оптимальной локально распределённой структуры [7]. Пользуясь этим принципом, авторы работы [8] создали две глубокие СНС, показывающие

результаты в 99.53% и 96.0% точности на базе лиц LFW.

В работе [9] исследуется задача распознавания лиц на фотографиях со значительными вариациями позы и положения лица, такими как фотографии пользователей из социальных сетей. Для распознавания лиц в таких условиях авторами предложена система Pose Invariant Person Recognition (PIPER), в которой ансамбль СНС используется на первом этапе для нахождения человеческой головы/лица, а затем происходит распознавание обнаруженных лиц при помощи системы DeepFace [10]. Для обучения и тестирования авторы работы собрали собственную базу изображений под названием PeopleInPhotoAlbums (PIPA), состоящую из **60000 тысяч** фотографий пользователей, на которой система PIPER показывает результат в 93.4% по сравнению с 89.3% у отдельно используемой DeepFace.

СНС могут применяться для снижения размерности пространства признаков. В работе [11] авторами представлена система под название FaceNet, в которой глубокая СНС используется для эффективного снижения размерности признаков, извлекаемых из изображения. Пространство признаков отображается в компактное евклидово пространство (один вектор занимает размер в 128 байт), в котором расстояние между векторами признаков изображений напрямую коррелирует с похожестью лиц на изображениях. Система достигает рекордной точности в 99.63% на базе лиц LFW.

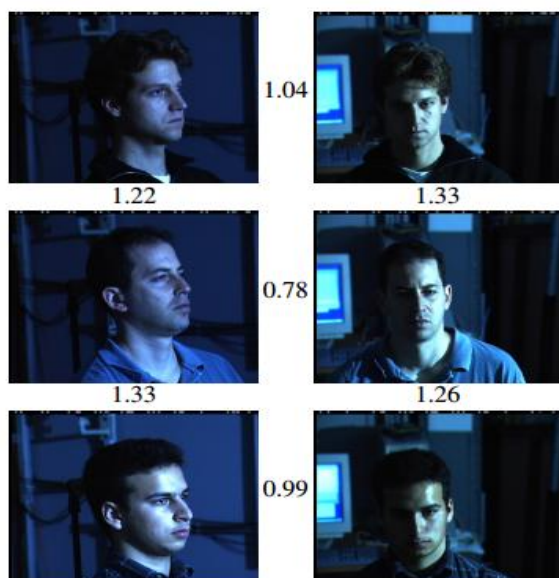


Рис.2. Расстояния в пространстве признаков между парами изображений после применения FaceNet. На основе порога в 1.1 можно судить о том, разные ли люди на изображениях [12]

Заключение

СНС могут применяться на различных этапах решения задачи распознавания лиц на изображениях:

1. В качестве основного классификатора или составной части каскада/ансамбля классификаторов при распознавании.

2. Для подготовки к распознаванию путём распознавания позы и положения тела и лица человека.

3. Для высокоэффективного выделения и снижения размерности признаков перед последующим распознаванием.

На основе приведённого краткого обзора, можно утверждать, что СНС являются одним из ключевых инструментов для решения задачи распознавания человеческих лиц на изображениях.

Литература

1. Hubel, D. H. Receptive fields and functional architecture of monkey striate cortex / D. H. Hubel, T. N. Wiesel // The Journal of physiology. – 1968. – no. 195(1). – Pp. 215–243.

2. Krizhevsky, A. Imagenet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in neural information processing systems. – 2012. – Pp. 1097–1105.

3. Bundy, A. Difference of Gaussians / A. Bundy, L. Wallen // Catalogue of Artificial Intelligence Tools. – 1984. – Pp. 30–35.

4. Parkhi, O. M. Deep Face Recognition / O. M. Parkhi, A. Vedaldi, A. Zisserman // British Machine Vision Conference. – 2015.

5. Huang, G. B. Labeled faces in the wild: A database for studying face recognition in unconstrained environments / G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller // Technical Report 07-49, University of Massachusetts, Amherst. – October 2007.

6. Szegedy, C. Going deeper with convolutions / C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich // Technical report, arXiv:1409.4842. – 2014.

7. Arora S. Provable Bounds for Learning Some Deep Representations / A. Bhaskara, R. Ge, T. Ma // arXiv:1310.6343v1. – 2013.

8. Sun, Y. DeepID3: Face Recognition with Very Deep Neural Networks / Y. Sun, D. Liang, X. Wang, X. Tang // arXiv:1502.00873v1. – 2015.

9. Zhang, N. Beyond Frontal Faces: Improving Person Recognition Using Multiple Cues / N. Zhang, M. Paluri, Y. Taigman, R. Fergus, L. Bourdev // arXiv:1501.05703v2. – 2013.

10. Taigman, Y. DeepFace: Closing the Gap to Human-Level Performance in Face Verification / Y. Taigman, M. Yang, M. Ranzato, L. Wolf // Conference on Computer Vision and Pattern Recognition (CVPR). – June 24, 2014.

11. Schroff, F. FaceNet: A Unified Embedding for Face Recognition and Clustering / F. Schroff, D. Kalenichenko, J. Philbin // arXiv:1503.03832v3. – 2015.