



# VISUAL PRODUCT RECOMMENDATION USING NEURAL AGGREGATION

## GOAL

Разработка прототипа мобильного приложения для классификации интересов пользователя по изображениям в галерее его мобильного устройства на основании сверточных нейронных сетей с применением нейронного блока агрегации с механизмом внимания, ранее успешно применявшимся для решения задачи классификации видео

## CONTRIBUTION

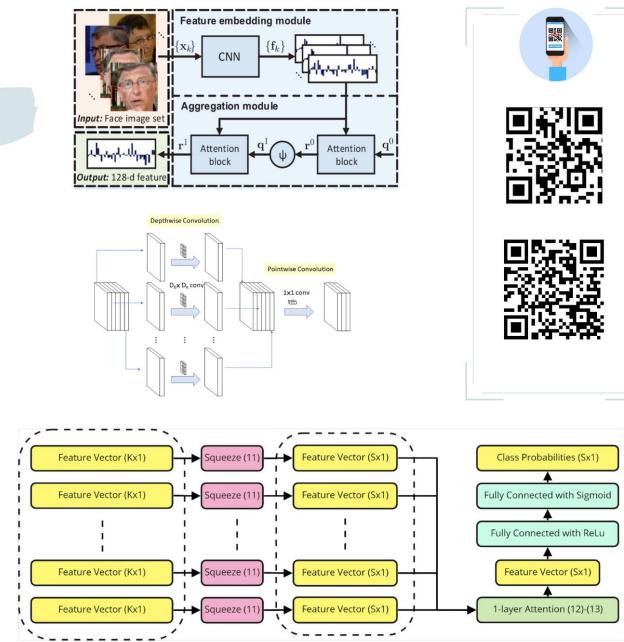
Разработан алгоритм классификации интересов пользователя по набору изображений на основе модификаций нейроагрегационного модуля. Предложенный алгоритм реализован в виде прототипа программной системы на мобильной платформе Android.

## MOTIVATION



Традиционные рекомендательные системы обладают существенным недостатком: они не используют визуальную информацию о предложениях, которые заинтересовали пользователя.

## APPROACH



## AGGREGATION

$$s_n(m) = W_s \mathbf{x}_n(m)$$

$$w(\mathbf{x}_n(m)) = \frac{\exp(\mathbf{q}^T s_n(m))}{\sum_{j=1}^{M_n} \exp(\mathbf{q}^T s_n(j))}$$

$$\mathbf{x}_n = \sum_{m=1}^{M_n} w(\mathbf{x}_n(m)) \mathbf{s}_n(m)$$

$$\mathbf{q}^1 = W_e \tanh(W \mathbf{x}_n + b)$$

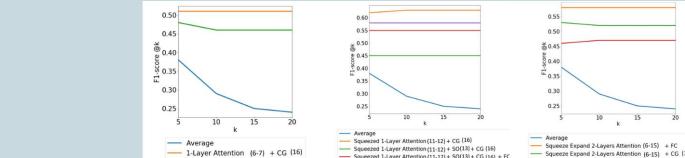
## RESULTS

- Проведены экспериментальные исследования по классификации наборов изображений на открытых данных Amazon Fashion.
- Показано, что предложенный подход работает в два раза лучше ( $F1@10=0.58$  для предложенного подхода против  $F1@10=0.29$  для традиционного подхода) чем традиционное усреднение векторов признаков.
- Было установлено, что алгоритм сжатия пространства признаков перед подсчетом весов агрегации не ведет к значительным потерям в качестве, позволяя при этом сократить количество обучаемых параметров и, следовательно, объем результирующей модели в 2 раза (781 000 обучаемых параметров для предложенного подхода против 1 570 024 обучаемых параметра для подхода без сжатия пространства признаков).

Метод агрегации	Число параметров	Размер (Мб)	Скорость принятия решения (мс)
Average			0.66 0.73
1-Layer Attention (6)-(7) + CG (16) + FC (16) + FC	1 025 000	4	0.59 0.8
Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	1 570 024	6	0.78 0.93
Squeeze Expand 2-Layers Attention (6)-(15) + FC		43	0.77 0.88
Average			0.77 0.9
1-Layer Attention (6)-(7) + CG (16) + FC			0.93 0.9
Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC			0.91 0.83
Squeeze Expand 2-Layers Attention (6)-(15) + FC			0.9 0.86
Average			0.68 0.71
1-Layer Attention (6)-(7) + CG (16) + FC			0.93 0.9
Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	781 800	3	0.94 0.82
Squeeze Expand 2-Layers Attention (6)-(15) + FC	1 109 736	4.5	0.93 0.84
Average			0.69 0.7
1-Layer Attention (6)-(7) + CG (16) + FC			0.94 0.89
Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC			0.95 0.81
Squeeze Expand 2-Layers Attention (6)-(15) + FC			0.94 0.84

Размер получившейся нейросетевой модели составляет всего 3 Мб (16 Мб вместе с предобученной CHC MobileNet v1)

k	Метод агрегации	Recall@k	Precision@k
5	Average	0.59	0.8
	1-Layer Attention (6)-(7) + CG (16) + FC	0.78	0.93
	Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	0.77	0.88
	Squeeze Expand 2-Layers Attention (6)-(15) + FC	0.77	0.9
10	Average	0.66	0.73
	1-Layer Attention (6)-(7) + CG (16) + FC	0.93	0.9
	Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	0.91	0.83
	Squeeze Expand 2-Layers Attention (6)-(15) + FC	0.9	0.86
15	Average	0.68	0.71
	1-Layer Attention (6)-(7) + CG (16) + FC	0.93	0.9
	Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	0.94	0.82
	Squeeze Expand 2-Layers Attention (6)-(15) + FC	0.93	0.84
20	Average	0.69	0.7
	1-Layer Attention (6)-(7) + CG (16) + FC	0.94	0.89
	Squeeze 1-Layer Attention (11)-(12) + SO (3) + FC	0.95	0.81
	Squeeze Expand 2-Layers Attention (6)-(15) + FC	0.94	0.84



Агрегация с помощью одного слоя сжатия с линейным оператором сжатия пространства признаков достигается  $F1$  мера равная 0.58, в два раза превышающая  $F1$  меру традиционной классификации среднего вектора признаков

