

ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКОЙ МОДЕЛИ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

Фуркатбек Рейимбергенович НУРЖАНОВ

доцент

Академия Вооруженных Сил Республики Узбекистан
Ташкент, Узбекистан

Асрор Махмадустанович БОЙТЕМИРОВ

докторант

Ташкентский университет информационных технологий имени Мухаммада ал-Хоразмий
Ташкент, Узбекистан

Элёр Нормуродович МУМИНОВ

ассистент

Ташкентский университет информационных технологий имени Мухаммада ал-Хоразмий
Ташкент, Узбекистан

Аннотация

В статье рассмотрены основные классы задач распознавания человека по изображению лица. Приведены описание и результаты предварительных экспериментов по созданию системы контроля доступа на основе анализа изображения лица человека. Указаны преимущества и перспективы использования математических моделей нейронных сетей для распознавания лиц.

Ключевые слова: нейронные сети, распознавание изображений лиц, методы, математическая модель, алгоритм.

YUZNI TANISHDA NEYRON TARMOQLARNING MATEMATIK MODELINI QO'LLASH

Furkatbek Reyimberganovich NURJANOV

dotsent

O'zbekiston Respublikasi Qurolli Kuchlari Akademiyasi
Toshkent, O'zbekiston

Boytemirov Asror Maxmado'stovich

doktorant

Muhammad al-Xorazmiy nomidagi
Toshkent axborot texnologiyalari Universiteti
Toshkent, O'zbekiston

asrorxon1007@gmail.com

Muminov Elyor Normurodovich

assistent

Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari
Universiteti
Toshkent, O'zbekiston

Annotatsiya

Ushbu maqolada yuz tasvirini aniqlashda neyron tarmoqlar muammosini hal qilish fan sohasida tobora dolzarb bo'lib bormoqda. Shu munosabat bilan yuz tasvirini aniqlashga asoslangan neyron tarmoqlarning matematik modellaridan foydalanish usullari va afzalliklari ko'rib chiqiladi. Bundan tashqari, tavsiya etilgan matematik model va neyron tarmoq algoritmi yuz tasvirini aniqlash ilovalari uchun juda mos keladi. Maqolada yuzni aniqlash maqsadida neyron tarmoqlarga asoslangan matematik modeldan amaliy foydalanish muhokama qilinadi.

Tayanch so'zlar: neyron tarmoqlari, yuz tasvirini aniqlash, usullar, matematik model, algoritm.

В мире особое внимание уделяется изучению сложных объектов в процессе развития биометрических технологий, а также совершенствованию и разработке систем идентификации и распознавания личности по лицу на основе обработки изображения с помощью современных компьютерных технологий. В этой связи существенно повышаются требования к системам распознавания и идентификации личности на основе биометрических технологий, которые широко используются при контроле в аэропортах и метрополитене, доступе в здание или программу, видеонаблюдении, в криминалистике и ряде других сфер.

Распознавание человека по изображению лица выделяется среди биометрических систем тем [5], что, во-первых, не требует специального дорогостоящего оборудования – для большинства приложений достаточно персонального компьютера и обычной видеокамеры. Во-вторых, отсутствует физический контакт человека с устройствами – не нужно ни к чему прикасаться или специально останавливаться и ждать срабатывания системы.

К недостаткам распознавания человека по изображению лица следует отнести то, что сама по себе такая система не обеспечивает стопроцентной надежности идентификации. Там, где требуется высокая надежность, применяют комбинирование нескольких биометрических методов.

Задача обнаружения лица на изображении является первым шагом, предобработкой в процессе идентификации личности человека по изображению (например, узнавания лица, распознавания выражения лица).

Эта задача считается более чем простой для человеческого зрения, однако при попытке построения автоматической системы обнаружения лиц приходится столкнуться со следующими сложностями:

сильно варьирующийся внешний вид лица у разных людей;

даже относительно небольшое изменение ориентации лица относительно камеры влечет за собой серьезное изменение изображения лица;

присутствие индивидуальных особенностей (усы, борода, очки, морщины и т. д.) существенно осложняет автоматическое распознавание;

изменение выражения лица может сильно сказаться на том, как лицо выглядит на изображении;

часть лица может быть невидима (закрыта другими предметами) на изображении;

условия съемки (освещение, цветовой баланс камеры, искажения изображения, привносимые оптикой системы, качество изображения) в значительной степени влияют на получающееся изображение лица [3].

В настоящее время для решения задачи обнаружения лица используются следующие подходы [5]:

- метод главных компонент;
- факторный анализ;
- моментный анализ;
- линейный дискриминантный анализ;
- метод опорных векторов;
- скрытые марковские модели;
- active appearance models (активные модели внешнего вида);
- вейвлетный анализ;
- подходы, основанные на искусственных нейронных сетях [1], [4], [6], [7], [8], [9].

В связи с этим в последнее время интенсивно проводятся исследования по созданию систем искусственного интеллекта. Их основной целью является

разработка математических и программных средств моделирования процессов человеческого мышления для автоматического решения различных прикладных и теоретических задач.

Поэтому и возникает задача увеличения точности распознавания, а значит, и необходимость использования новых подходов и алгоритмов обработки изображений. Чтобы качественно улучшить систему распознавания, рассмотрим применение нейронной сети.

Нейросеть – это математическая модель в виде программного и аппаратного воплощения, строящаяся на принципах функционирования биологических нейросетей. Сегодня такие сети активно используют в практических целях за счет возможности не только разработки, но и обучения. Их применяют для прогнозирования, распознавания образов, машинного перевода, распознавания аудио и т.д.

Нейросетевые методы предлагают иной подход к решению задачи распознавания образов. Исходя из этого применение нейронных сетей для задачи распознавания человека по изображению лица является перспективным направлением.

Несмотря на большое разнообразие вариантов нейронных сетей, все они имеют общие черты. Так, все они, так же, как и мозг человека, состоят из большого числа связанных между собой однотипных элементов – *нейронов*, которые имитируют нейроны головного мозга. На рис. 1 показана схема нейрона.



Рис. 1. Схема нейрона

Из рисунка видно, что искусственный нейрон, так же, как и живой, состоит из синапсов, связывающих входы нейрона с ядром; ядра нейрона, которое осуществляет обработку входных сигналов и аксона, который связывает нейрон с нейронами следующего слоя. Каждый синапс имеет вес, который определяет, насколько соответствующий вход нейрона влияет на его состояние. Состояние нейрона определяется по формуле:

$$S = \sum_{i=1}^n x_i * w_i \quad (1)$$

$$S = \sum_{k=1}^N k^2, \quad (1.1)$$

где

n – число входов нейрона

x_i – значение i -го входа нейрона

w_i – вес i -го синапса

Затем определяется значение аксона нейрона по формуле:

$$Y = f(S), \quad (2)$$

где f – некоторая функция, которая называется *активационной*. Наиболее часто в качестве активационной функции используется так называемый *сигмоид*.

В общем случае задача обучения НС сводится к нахождению некой функциональной зависимости $Y=F(X)$, где X – входной, а Y – выходной векторы. В общем случае такая задача при ограниченном наборе входных данных имеет бесконечное множество решений. Для ограничения пространства поиска при обучении ставится задача минимизации целевой функции ошибки НС, которая находится по методу наименьших квадратов.

Обучение нейросети производится методом градиентного спуска, на каждой итерации изменение веса производится по формуле:

$$E(w) = \frac{1}{2} \sum_{j=1}^p (y_i - d_i)^2 (3),$$

где

y_i – значение j -го выхода нейросети

d_i – целевое значение j -го выхода

p – число нейронов в выходном слое

Обучение нейросети производится методом градиентного спуска, т.е. на каждой итерации изменение веса производится по формуле:

$$\Delta w_{ij} = -\mu \cdot \frac{dE}{dw_{ij}} \quad (4),$$

где h – параметр, определяющий скорость обучения.

$$\frac{dE}{dw_{ij}} = \frac{dE}{dy_{ij}} \cdot \frac{dy_i}{dS_j} \cdot \frac{dS_i}{dw_{ij}}, \quad (5)$$

где

y_i – значение выхода j -го нейрона

S_j – взвешенная сумма входных сигналов, определяемая по формуле (1)

При этом множитель

$$\frac{dS_i}{dw_{ij}} = x_i, \quad (6)$$

где

x_i – значение i -го входа нейрона.

Далее рассмотрим определение первого множителя формулы (5)

$$\frac{dE}{dy_{ij}} = \sum k \frac{dE}{dy_k} \cdot \frac{dy_k}{dS_k} \cdot \frac{dS_k}{dy_i} = \sum k \frac{dE}{dy_k} \cdot \frac{dy_k}{dS_k} \cdot w_{jk}^{(n+1)}, \quad (7)$$

где

k – число нейронов в слое $n+1$.

Введем вспомогательную переменную:

$$\delta_j^{(n)} = \frac{dE}{dy_i} \cdot \frac{dy_i}{dS_j} \quad (8)$$

Тогда мы сможем определить рекурсивную формулу для определения n -ного слоя, если нам известно следующего $(n+1)$ -го слоя.

$$\delta_j^{(n)} = \left[\sum k \delta_k^{(n+1)} \cdot w_{jk}^{(n+1)} \right] \cdot \frac{dy_i}{dS_j}, \quad (9)$$

Нахождение же для последнего слоя НС не представляет трудности, так как нам известен целевой вектор, т.е. вектор тех значений, которые должна выдавать НС при данном наборе входных значений [8].

$$\delta_j^{(N)} = \left(y_i^{(N)} - d_i \right) \cdot \frac{dy_i}{ds_j}, \quad (10)$$

И наконец запишем формулу (4) в раскрытом виде:

$$\Delta\omega_{ij}^n = -\mu \cdot \delta_j^{(n)} \cdot x_i^n, \quad (11)$$

Рассмотрим полный алгоритм обучения нейросети:

1. Подать на вход НС один из требуемых образов и определить значения выходов нейронов нейросети.

2. Рассчитать для выходного слоя НС по формуле (10) и рассчитать изменения весов выходного слоя N по формуле (11).

3. Рассчитать по формулам (9) и (11) соответственно и $\Delta\omega_{ij}^{(N)}$ для остальных слоев нейронных сетей, $n = N - 1 \dots 1$

4. Скорректировать все веса НС

$$\Delta\omega_{ij}^{(n)} t = \omega_{ij}^{(n)}(t - 1) + \Delta\omega_{ij}^{(n)}(t),$$

5. Если ошибка существенна, то перейти на шаг 1

Простейший метод градиентного спуска, рассмотренный выше, очень неэффективен в случае, когда производные по различным весам сильно отличаются. Это соответствует ситуации, когда значение функции S для некоторых нейронов близко по модулю к 1 или когда модуль некоторых весов много больше 1. В этом случае для плавного уменьшения ошибки надо выбирать очень маленькую скорость обучения, но при этом обучение может занять nepозволительно много времени [1].

Простейшим методом усовершенствования градиентного спуска является введение *момента t*, когда влияние градиента на изменение весов меняется со временем. Тогда формула (11.1) примет вид:

$$\Delta\omega_{ij}^{(n)} t = -\mu \cdot \delta_j^{(n)} \cdot x_i^n + \mu\Delta\omega_{ij}^n(t - 1), \quad (11.1)$$

Дополнительным преимуществом от введения *момента* является способность алгоритма преодолевать мелкие локальные минимумы.

В связи с этим на этапе обучения нейронной сети предъявляется предварительно подготовленный набор изображений определенного размера, содержащий как изображения лица, так и нелицевые изображения (фон, части лиц и т.п.).

Прямое сравнение признаков применялось совместно с корреляционным методом сравнения изображений с выборкой. Анализируется входное множество признаков $\{F\} = \{C_{i,j}^{\max}, i = \overline{1, N}, j = \overline{1, M}\}$, где N – число классов, M – число образцов класса в обучающей выборке. Были рассмотрены следующие варианты классификации [8], [9]:

1. По максимуму корреляции. Результат обнаружения вычисляется как:

$$\{P_i\} = \{C_i^{\max}\} = \left\{ \max_{j=\overline{1, M}} C_{i,j}^{\max}, i = \overline{1, N} \right\} \quad (12)$$

2. По числу превышений фиксированного порога по всей выборке. Введем некоторый порог $T \in [0,1]$. Определим результат обнаружения по следующей формуле:

$$\{P_i\} = \left\{ \sum_{j=1}^M \begin{cases} 1, C_{i,j}^{\max} \geq T \\ 0, C_{i,j}^{\max} < T \end{cases}, i = \overline{1, N} \right\} \quad (13)$$

3. Интегральный критерий с фиксированным порогом. Аналогично вводим порог $T \in [0,1]$. Результат обнаружения вычислим по формуле:

$$\{P_i\} = \left\{ \sum_{j=1}^M \begin{cases} C_{i,j}^{\max} - T, C_{i,j}^{\max} \geq T \\ 0, C_{i,j}^{\max} < T \end{cases}, i = \overline{1, N} \right\} \quad (15)$$

4. Интегральный критерий с плавающим порогом. Введем два порога $T_{\min \in [0,1]}, T_{\max \in [0,1]}$. Вычислим максимум корреляции по всему множеству C^{\max} , если $C^{\max} < T_{\min}$, то полагаем результат классификации гарантированно неудачным и полагаем $\{P_i\} = \{0, i = \overline{1, N}\}$. В противном случае вычисляем конечный порог $T = C^{\max} \cdot T_{\max}$ и результат обнаружения получим по формуле (14).

5. Комплексный критерий. Вводим дополнительные пороги $T_{abs} \in [0,1], T_{mid} \in [0,1]$, так что $T_{min} < T_{mid} < T_{max} < T_{abs}$. Если $C^{max} < T_{min}$, то полагаем результат классификации гарантированно неудачным и полагаем $\{P_i\} = \{0, 1 = \overline{1, N}\}$. В противном случае применяем следующий алгоритм:

Если $C^{max} > T_{abs}$, то вычислим результат по критерию 1, иначе если $C^{max} > T_{max}$, то вычисляем результат по критерию 2, приняв $T = T_{max}$, иначе если $C^{max} > T_{mid}$, то вычислим результат по критерию 3, приняв $T = T_{mid}$ иначе если $C^{max} > T_{min}$, то вычисляем результат по критерию 3, приняв $T = T_{min}$.

6. Комплексный критерий с усреднением. Вычисляем минимум корреляции по всему множеству C^{min} . Если $C^{max} > T_{min}$, то полагаем результат классификации гарантированно неудачным и полагаем $\{P_i\} = \{0, 1 = \overline{1, N}\}$. В противном случае применяем следующий алгоритм и вычислим частичные результаты для каждого из подходящих условий:

Если $C^{max} > T_{abs}$, то вычисляем частичный результат по критерию 1. Если $C^{max} > T_{max}$, то вычислим результат по критерию 2, приняв $T = T_{max}$. Если $C^{max} > T_{mid}$, то вычислим результат по критерию 4, приняв $T_{max} = T_{mid}$ в критерии. Если $C^{max} > T_{min}$, то вычисляем результат по критерию 3, приняв $T = C^{min}$. За конечное $\{P_i\}$ принимается среднее частичных результатов по числу выполненных условий. Для упрощения комплексирования результатов распознавания различных классификаторов $\{P_i\}$ [8], [9] масштабируется в диапазон $[0,1]$ по следующей формуле:

$$\{P_i^S\} = \left\{ \frac{P_i - P^{min}}{P^{max} - P^{min}}, i = \overline{1, N} \right\}, \quad (15)$$

где $P^{max} = \max_{j=\overline{1, N}} P_j, P^{min} = \min_{j=\overline{1, N}} P_j$

База ORL содержит 400 изображений: 40 человек по 10 изображений каждого. Для каждого эксперимента база делилась случайным образом на две части, обучающую и тестовую, по пять изображений одного человека в

каждой части. Для обучения сети и сравнения с неизвестным изображением использовалась одна и та же обучающая выборка.

Исследовались возможности распознавания на основе полученных главных компонентов и возможности реконструкции изображения в зависимости от следующих факторов:

- количество обучающих циклов;
- число скрытых нейронов;
- разрешение изображения: ORL/1 (92x112 пикселей, исходный размер), ORL/2 (46x56), ORL/4 (23x28);
- различная случайная разбивка на тестовую и тренировочную части.

Программное обеспечение (ПО) поддерживает источники видео разрешением только 640x480 пикселей. Используемые в качестве источника видеокamеры должны поддерживать DirectShow и возможность установки режима RGB 640x480x24 или RGB 640x480x32.

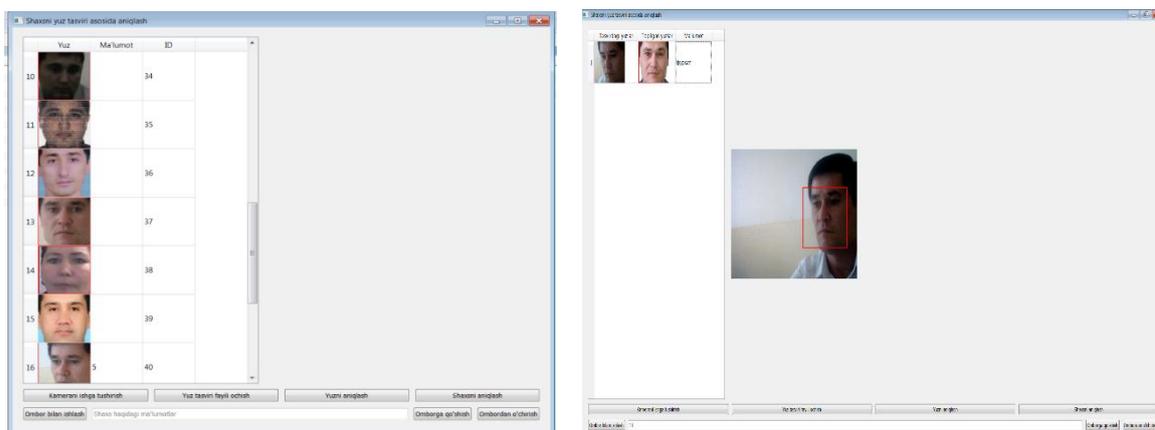


Рис. 2. Программный продукт – интерфейс лица личности

Сформирована обучающая выборка из 110 изображений лиц 5 человек (5x22 изображений лиц). Каждое изображение подверглось предобработке – переводу в градации серого, выравниванию гистограмм, контрастированию и масштабированию до размера 128x128 пикселей для алгоритма и 48x48, 32x32 пикселя для различных зон интереса метода.

Полученные результаты можно охарактеризовать следующим образом: ошибка реконструкции быстро уменьшается в течение первых 10-20 шагов и дальше практически не изменяется, ошибка распознавания имеет похожую

тенденцию. Обе эти величины слегка колеблются вследствие случайного порядка обучающих образов.

Тестирование проводилось по 6 зонам в области лица: прямоугольная область глаз, отдельные зоны для каждого глаза, носа, рта и целое изображение лица. Подобные тесты с аналогичными результатами были проведены для разложений по существенно большему и меньшему числу собственных векторов. Это позволило сделать вывод о малой пригодности данного метода для построения признаков в реальных условиях наличия геометрических искажений и нестабильности освещения идентификации лиц, несмотря на высокое быстродействие. Для лиц, присутствующих в обучающей выборке, результаты представлены в таблице 1.

Результаты распознавания для пользователей, представленных в базе данных

Таблица 1.

Имя пользователя	Число срабатываний	Число отказов в доступе	Число ошибок распознавания	Процент успешной идентификации
Александр	193	1	0	90.5
Бултаков	187	0	0	88.3
Фуркат	175	41	0	92.8
Хуршид	112	9	0	76.6
Сардор	91	9	0	90.3

Процент успешной идентификации: 92.8%

Ошибок распознавания: 0.0%

Отказов в доступе: 7.2%

Здесь: число срабатываний – это количество раз, когда система выдала информацию об успешной идентификации пользователя, успешном распознавании личности; число отказов в доступе – количество кадров, лицо на которых было обнаружено, но не могло быть отнесено ни к одному из классов; число ошибок – количество раз, когда пользователь был ошибочно опознан как другой человек из базы данных.

В таблице 2 представлены результаты идентификации для посторонних лиц.

Результаты распознавания для лиц, отсутствующих в базе данных

Таблица 2.

Имя	Число срабатываний	Число ошибок по статистике	Число отказов в доступе	Процент ошибок
Владимир	0	0	61	0
Елена	0	0	173	0
Виталик	0	0	55	0
Михаил	10	0	205	4.7

Ошибок распознавания: 4.7%

Ошибочных разрешений доступа: 0.0%

Здесь: число срабатываний – количество раз ошибочного обнаружения постороннего как пользователя из базы данных; число ошибок по статистике – количество раз, когда пользователь был ошибочно допущен к системе как другой человек из базы данных (число ошибочных обнаружений достигло требуемого для принятия решений числа накоплений результатов); число отказов в доступе – количество кадров, лицо на которых было обнаружено, но не могло быть отнесено ни к одному из классов.

По результатам можно видеть, что, несмотря на имеющиеся ошибочные срабатывания, за счет статистической обработки доступ к системе был закрыт во всех случаях. Исходный материал – растровое изображение (линейный размер изображения лица – не меньше 150x100 пикселей, не более 1000x800 пикселей). Масштабные преобразования относительно эталона – уменьшение/увеличение в 6 раз. Вариация освещения – в пределах 10 – 15% от среднего уровня яркости эталонного изображения. Описание инвариантно к фону. Поддержка многоэталонности (борода/без бороды, в очках/без очков).

Вероятность появления искомой фотографии в первой пятерке рейтинга (при базе в 500 эталонов) ~ 97%. Время формирования описания ~ 1 с. Время сравнения ~ 3 с (для базы в 5000 человек). Web-интерфейс (защищенный паролем) – есть.

В заключение можно отметить, что впервые на русском языке дан обзор нейросетевых методов распознавания человека по изображению лица, указаны их преимущества и перспективы. В работе представлены математические модели постобработки изображения лица, рассматривается их применение к задаче выделения ключевых точек на лице личности. Определены этапы проектирования разработки нового алгоритма идентификации изображений лиц в процессе формирования математической модели алгоритма идентификации изображений лица личности. Приведены описание и результаты предварительных экспериментов по созданию системы контроля доступа на основе анализа изображения лица человека.

Практическая значимость исследования заключается в том, что предложенные программные алгоритмы могут быть использованы для проектирования систем обработки изображений лица и деталей, которые являются компонентом современных систем моделирования для идентификации и сравнения идентификаторов распознавания лиц.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА:

1. Алексеева В.И., Коробка Ж.И. Искусственные нейронные сети: рекомендательный библиографический указатель. – Краматорск: ДГМА, 2014. – С. 25.
2. Друки А.А. Система поиска, выделения и распознавания лиц на изображениях // Средства и системы обработки и анализа данных, 2011. – С.16–25.
3. Зайнидинов Х.Н., Нуржанов Ф.Р., Даулетов А.Ю. Параллельная обработка изображения лица личности //Вестник Туринского политехнического университета в городе Ташкенте, 2018. – № 4. – С. 28-33.

4. Потапов А.С. Искусственный интеллект и универсальное мышление. – СПб.: Политехника, 2012. – 711 с.
5. Самаль Д.И., Старовойтов В.В. Подходы и методы распознавания людей по фотопортретам. – Минск: Ин-т. техн. кибернетики НАН Беларуси, 1998. – 54 с.
6. Скурихин А.Н. Генетические алгоритмы //Новости искусственного интеллекта, 1995. – №4. – С. 6-46.
7. Gutta S., Wechsler H. Face recognition using hybrid classifiers // Pattern Recognition. – 1997. – Vol. 30. – P. 539-553.
8. Nurjanov F.R. Mathematical models image face of the personality IJARSET: International journal of Advanced Research in Science Engineering and Technology Volume 6 Issue 7, July 2019, India 06/07/2019, 9 pages.
9. Nurjanov F.R. Method and algorithm for identifying the parameters of the image face person // International Journal of Science and Research (IJSR). Volume 8 Issue 6, June 2019, India 08/06/2019. 7 pages. Scientific Journal Impact Factor, № 23; Impact Factor 7,426.
10. Sadykhov R., Vatkin M. New training algorithm of neural network “Neokognitron” for a recognition of hand-written symbols // Pattern Recognition and Image Processing. – 2001. – P. 203-212