

Аликрицкий Михаил Сергеевич, магистрант,
Поволжский государственный университет
телекоммуникаций и информатики

КЛАССИФИКАЦИЯ ЭМОЦИЙ ПО ИЗОБРАЖЕНИЯМ ЛИЦ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ КЛАССИЧЕСКИХ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ И СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Аннотация. В работе рассматривается задача автоматической классификации эмоций по изображениям лиц. Целью исследования является анализ и экспериментальное сравнение применимости классических методов компьютерного зрения и методов глубокого обучения для распознавания эмоциональных состояний. Полученные результаты демонстрируют преимущества методов глубокого обучения при решении данной задачи.

Ключевые слова: Эмоции, компьютерное зрение, классификация, SVM, сверточные сети.

Автоматическая классификация эмоций по изображениям лиц является одной из актуальных задач в области компьютерного зрения и анализа изображений [1]. Распознавание эмоциональных состояний человека находит применение в системах человеко-машинного взаимодействия, интеллектуальных помощниках, анализе поведения пользователей, а также в медицинских и образовательных технологиях. Несмотря на значительный прогресс в данной области, задача остается сложной из-за вариативности выражений лиц, различий в освещении, позах и индивидуальных особенностях людей, а также субъективности интерпретации эмоциональных состояний [1, 6].

В рамках данной работы рассматриваются два основных подхода к решению задачи классификации эмоций: классические методы компьютерного зрения, основанные на извлечении ручных признаков, и методы глубокого обучения, использующие сверточные нейронные сети [2, 5]. В качестве дата-сета был использован FER2013. Основная цель исследования заключается в проверке возможности эффективной классификации эмоций с помощью указанных подходов, а также в их сравнительном анализе по качеству распознавания.

В качестве базового классического подхода был рассмотрен метод опорных векторов, использующий локальные бинарные шаблоны в качестве признаков. Локальные бинарные шаблоны широко применяются для анализа текстур и описания локальной структуры изображения и доказали свою эффективность в задачах распознавания лиц и выражений при ограниченных вычислительных ресурсах [3, 7]. Для каждого изображения лица вычислялось гистограммное представление LBP-признаков, которое затем подавалось на вход классификатору. Результаты классификации для данного подхода представлены на рисунке 1.



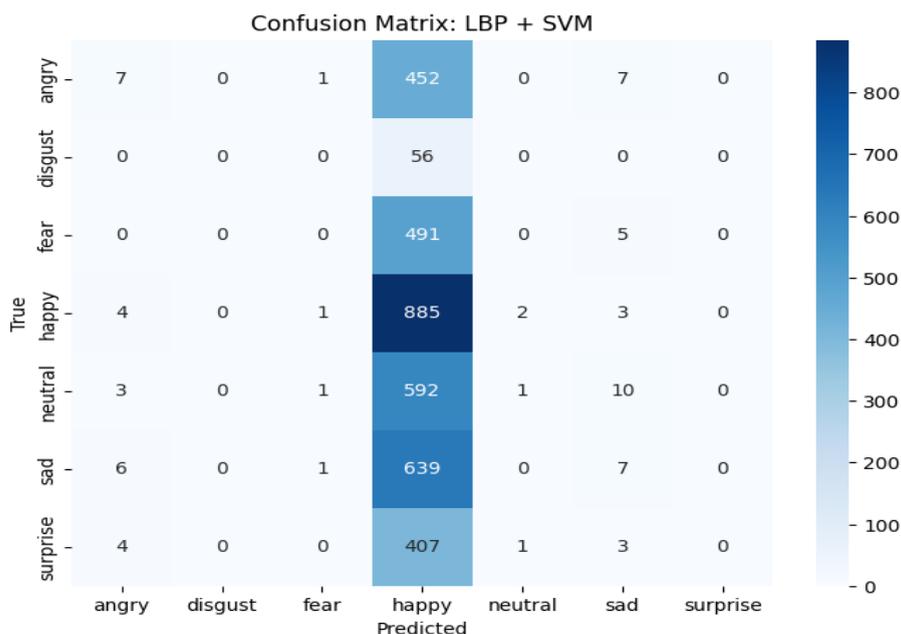


Рисунок 1. Матрица ошибок для метода LBP + SVM

Анализ полученных результатов показывает, что использование LBP-признаков в сочетании с SVM позволяет различать эмоциональные классы лишь на базовом уровне. Итоговая точность классификации составила 0,25, что

указывает на ограниченную выразительность ручных текстурных признаков при решении задачи распознавания эмоций по изображениям лиц.

В качестве второго классического подхода был реализован метод, основанный на использовании признаков ориентированных градиентов. Признаки HOG описывают форму объектов за счет анализа направлений градиентов яркости и широко применяются в задачах распознавания образов, включая детекцию лиц и анализ мимики [2, 8]. После извлечения HOG-признаков для всех изображений обучение классификатора SVM проводилось с линейным ядром. Матрица ошибок для данного метода представлена на рисунке 2.

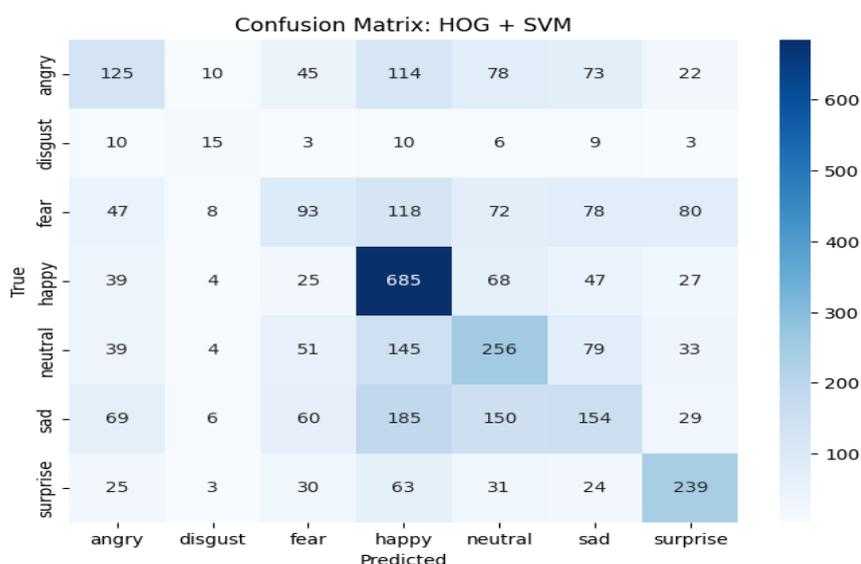


Рисунок 2. Матрица ошибок для метода HOG + SVM



По результатам эксперимента было установлено, что использование признаков HOG позволяет значительно повысить качество классификации по сравнению с LBP. Точность классификации составила 0,44, что свидетельствует о более высокой информативности признаков формы для анализа мимических выражений. Тем не менее, данный подход также демонстрирует ограничения, связанные с необходимостью ручного подбора признаков и невозможностью адаптации к сложным вариациям данных.

Для преодоления указанных ограничений был реализован метод глубокого обучения на основе сверточной нейронной сети. Сверточные нейронные сети позволяют автоматически извлекать иерархические признаки непосредственно из изображений, что делает их особенно эффективными для задач компьютерного зрения и распознавания эмоций в условиях высокой вариативности данных [5, 9]. Архитектура сети включала несколько сверточных и подвыборочных слоев, за которыми следовали полносвязные слои и слой классификации.

Обучение нейронной сети проводилось на обучающей выборке с использованием валидационной подвыборки для контроля процесса обучения и предотвращения переобучения. Оценка качества модели выполнялась на независимой тестовой выборке. Матрица ошибок для сверточной нейронной сети представлена на рисунке 3.

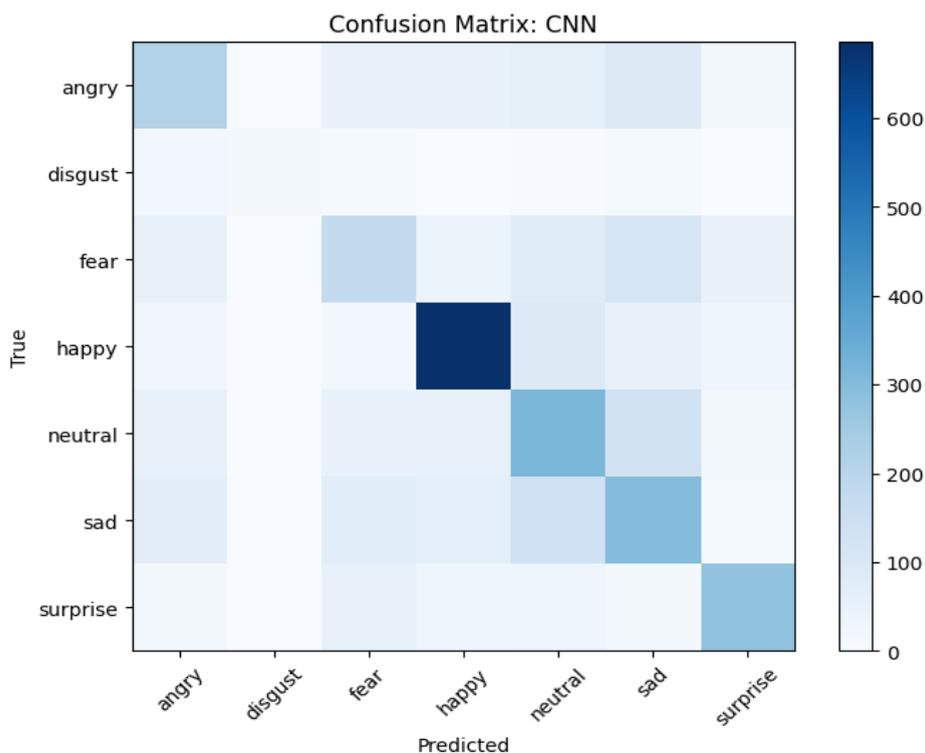


Рисунок 3. Матрица ошибок для сверточной нейронной сети

Экспериментальные результаты показали, что сверточная нейронная сеть превосходит классические методы по качеству классификации эмоций. Итоговая точность на тестовой выборке составила 0,55. Улучшение результатов объясняется способностью нейронной сети автоматически формировать высокоуровневые признаки, учитывающие сложные пространственные зависимости между элементами изображения.

Сравнительный анализ всех рассмотренных методов приведен в таблице 1.



Таблица 1.

Сравнение точности классификации различных методов

Метод	Признаки	Accuracy
SVM	LBP	0,25
SVM	HOG	0,44
CNN	Raw images	0,55

Проведенное исследование подтверждает, что задача классификации эмоций по изображениям лиц может быть решена как с использованием классических методов компьютерного зрения, так и с применением методов глубокого обучения.

Однако результаты экспериментов показывают явное преимущество сверточных нейронных сетей, обеспечивающих более высокую точность и лучшую способность к обобщению.

В заключение можно отметить, что классические методы остаются полезными в качестве базовых решений и инструментов для анализа, однако для построения высокоэффективных систем распознавания эмоций предпочтительным является использование методов глубокого обучения. В дальнейшем исследование может быть расширено за счет использования более сложных архитектур нейронных сетей, методов увеличения данных и анализа устойчивости моделей к различным искажениям изображений.

Список литературы:

1. Ekman P. Emotions Revealed: Recognizing Faces and Feelings / P. Ekman. – New York: Times Books, 2003. – 304 p.
2. Dalal N. Histograms of Oriented Gradients for Human Detection / N. Dalal, B. Triggs // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2005. – Vol. 1. – P. 886–893.
3. Ojala T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns / T. Ojala, M. Pietikäinen, T. Mäenpää // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24, No. 7. – P. 971–987.
4. Bishop C.M. Pattern Recognition and Machine Learning / C.M. Bishop. – New York: Springer, 2006. – 738 p.
5. Goodfellow I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – Cambridge: MIT Press, 2016. – 775 p.
6. Calvo R.A. Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications / R.A. Calvo, S. D’Mello // IEEE Transactions on Affective Computing. – 2010. – Vol. 1, No. 1. – P. 18–37.
7. Shan C. Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study / C. Shan, S. Gong, P.W. McOwan // Image and Vision Computing. – 2009. – Vol. 27, No. 6. – P. 803–816.
8. Happy S.L. Automatic Facial Expression Recognition Using Features of Salient Facial Patches / S.L. Happy, A. Routray // IEEE Transactions on Affective Computing. – 2015. – Vol. 6, No. 1. – P. 1–12.
9. Li S. Deep Facial Expression Recognition: A Survey / S. Li, W. Deng // IEEE Transactions on Affective Computing. – 2020. – Vol. 13, No. 3. – P. 1195–1215.

