

## Сравнение популярных библиотек компьютерного зрения для использования в приложении по распознаванию транзисторов

Макогон С.А., Зори С.А.  
Донецкий национальный технический университет  
makogon\_94@mail.ru, ik.ivt.rec@gmail.com

*Макогон С.А., Зори С.А. Сравнение популярных библиотек компьютерного зрения для использования в приложении по распознаванию транзисторов. В данной статье произведен обзор и сравнение библиотек OpenCV, AForge.NET, LTI и VXL для использования в приложениях по распознаванию образов и доказана целесообразность использования библиотеки OpenCV для приложений по распознаванию транзисторов.*

**Ключевые слова:** компьютерное зрение, распознавание образов, библиотека, OpenCV, AForge.NET, LTI, VXL, документация

### Введение

В настоящее время у разработчиков программного обеспечения существует большой интерес к отрасли компьютерного зрения. Это обусловлено тем, что данная технология может быть использована в разнообразных прикладных областях и гибко настроена под выполнение определенных типовых заданий.

Так как сообщество разработчиков постоянно развивается, на данный момент существует определенное количество библиотек для разных языков программирования, подстроенных для выполнения типовых классов задач. В статье произведен анализ и сравнение наиболее популярных на сегодняшний день библиотек компьютерного зрения и рассмотрена возможность их применения для решения практических задач по распознаванию транзисторов.

### Библиотека OpenCV

OpenCV (Open Source Computer Vision Library) представляет собой библиотеку программного обеспечения для компьютерного зрения с открытым исходным кодом и компьютерного обучения. OpenCV была создана для обеспечения общей инфраструктуры приложений для компьютерного зрения и ускорения использования восприятия машины в коммерческих продуктах. Будучи лицензированным BSD продуктом, OpenCV упрощает бизнес для использования и модификации кода [1].

Целью разработки данной библиотеки является повышение эффективности вычислений в приложениях реального времени. Язык C, на котором была написана библиотека, является оптимизированным. Библиотека OpenCV способна использовать многоядерные процессоры. В случае, если понадобится автоматическая оптимизация на различных аппаратных платформах Intel, возможно дополнительное приобретение библиотеки IPP, с английского Integrate Performance Primitives. В состав данной библиотеки входят процедуры с низкоуровневой оптимизацией, которые могут применяться для разнообразных алгоритмических областей. В спектр возможностей OpenCV входит автоматическое применение IPP во время выполнения какой-либо программы. На рис. 1 приведена диаграмма, сравнивающая быстродействие библиотек OpenCV, OpenCV + IPP, VXL и LTI [2].

В библиотеке более 2500 оптимизированных алгоритмов, которые включают в себя полный набор классических и современных алгоритмов компьютерного зрения и машинного обучения. Эти алгоритмы могут использоваться для обнаружения и распознавания лиц, идентификации объектов, классификации действий человека в видео, отслеживания движения камеры, отслеживания движущихся объектов, извлечения 3D-моделей объектов, создания 3D-облаков точек из стереокамер, сшивания изображений вместе для получения изображения высокого разрешения для всей сцены, найти похожие изображения из базы данных изображений, удалять красные глаза с изображений, сделанных с помощью вспышки, следить за движениями глаз, распознавать декорации и устанавливать маркеры, чтобы накладывать их на дополненную реальность и т. д. OpenCV имеет более 47 тысяч пользователей сообщества и предполагаемое количество загрузок, превышающих 14 миллионов. Библиотека широко используется в компаниях, исследовательских группах и правительственных органах.

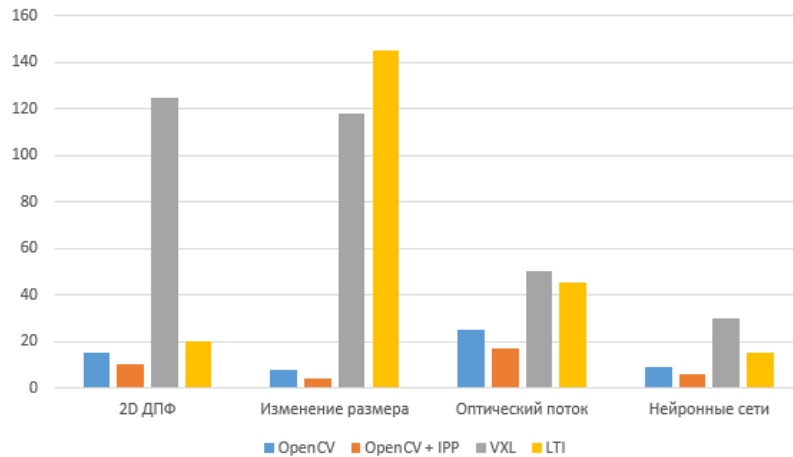


Рисунок 1 – Сравнение библиотек компьютерного зрения

Наряду с известными компаниями, такими как Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota, которые используют библиотеку, есть много стартапов, таких как Applied Minds, VideoSurf и Zeitera, которые широко используют OpenCV. Развернутые применения OpenCV охватывают диапазон от сшивания изображений уличного обзора вместе, обнаруживая вторжения в видео наблюдения в Израиле, контролируя оборудование шахт в Китае, помогая роботам перемещаться и собирать объекты в Willow Garage, обнаруживать несчастные случаи в бассейнах в Европе, запускать интерактивное искусство в Испании и Нью-Йорке, проверяя взлетно-посадочные полосы на наличие мусора в Турции, проверяя этикетки на продуктах на заводах по всему миру, чтобы быстро обнаружить лицо в Японии.

Библиотека имеет интерфейсы для C++, Python, Java и MATLAB и поддерживает Windows, Linux, Android и Mac OS. OpenCV ориентируется в основном на приложения в режиме реального времени и использует преимущества команд MMX и SSE, когда они доступны. В настоящее время активно развиваются полнофункциональные интерфейсы CUDA и OpenCL. Существует более 500 алгоритмов и около 10-кратного числа функций, которые составляют или поддерживают эти алгоритмы. OpenCV написан на языке C++ и имеет шаблонный интерфейс, который работает без проблем с контейнерами STL. На рис. 2 представлена общая структура проекта OpenCV.

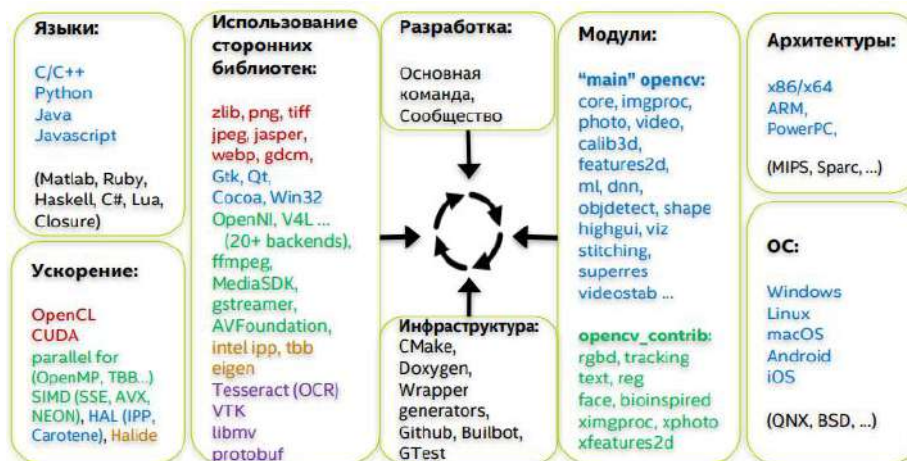


Рисунок 2 – Карта проекта OpenCV

### Библиотека AForge.NET

AForge.NET является библиотекой с открытым исходным кодом, созданной на языке C#, которая предназначена для разработчиков и исследователей в области компьютерного зрения. Кроме того, в библиотеке есть функционал для разработчиков в области искусственного интеллекта. Спектр возможностей библиотеки довольно широк: обработка изображений, нейронные сети, генетические алгоритмы, нечеткая логика, машинное обучение, робототехника и многое другое [3].

Библиотека включает несколько основных компонентов. AForge.Imaging — библиотека подпрограмм для обработки изображений и фильтров. AForge.Vision — библиотека компьютерного зрения. AForge.Video — набор библиотек для работы с видеoinформацией. AForge.Neuro — библиотека для выполнения разнообразных

действий и операций с нейронными сетями. AForge.Genetic — библиотека подпрограмм для использования генетических алгоритмов для решения различных задач. AForge.Fuzzy — библиотека для работы с нечеткой логикой. AForge.Robotics — библиотека, обеспечивающая поддержку некоторых методов, применяемых в сфере робототехники. AForge.MachineLearning — библиотека для работы с элементами машинного обучения [4] AForge.NET постоянно улучшается и прогрессирует. По данной библиотеке существует большое количество примеров, демонстрирующих ее работу, и html-документация, которая поддерживается в актуальном состоянии и помогает разработчикам в использовании данного фреймворка. Существует, так же как и у библиотеки OpenCV, активное сообщество, в котором можно подчеркнуть необходимую информацию, задавая вопросы разработчикам, или поделиться собственными наработками. Но, к сожалению, количество участников данного сообщества уступает своим количеством аналогичным по OpenCV. Еще одним ограничением на пути разработчика является тот факт, что вся документация по библиотеке написана только на английском языке. Ввиду этого возможны трудности в изучении и освоении данного фреймворка. На рис. 3 приведена общая структура библиотеки AForge.NET.

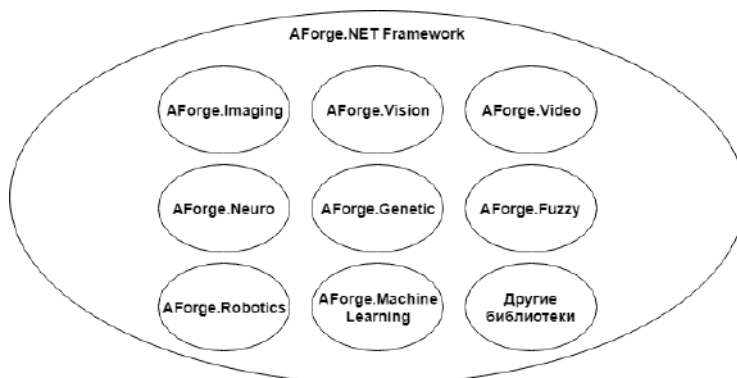


Рисунок 3 – Структура фреймворка AForge.NET

### **Библиотека VXL**

VXL — это набор библиотек, написанных на языке C++, которые предназначены для научных исследований и реализации технологий компьютерного зрения [5]. VXL была написана в ANSI/ISO C++ и предназначена для портативных платформ. Библиотека состоит из нескольких основных составляющих: VNL (числа) — численные алгоритмы и контейнеры, например, матрицы, векторы, оптимизаторы и т.д., VII (изображения) — загрузка, сохранение и редактирование изображений во многих наиболее распространенных форматах (также существует возможность работы с очень большими изображениями), VGL (геометрия) — геометрия точек, кривых и других элементарных объектов в одно-, двух- и трехмерном пространствах, VSL (входной и выходной потоки), VBL (основные шаблоны), VUL (утилиты) — разный функционал для независимых платформ. Кроме основных библиотек, входящих в состав VXL, существуют и дополнительные. Они отвечают за такие понятия, как численные алгоритмы, обработка изображений, системы координат, геометрия камеры, стерео, манипуляции с видеопотоком, восстановление структуры при движении камеры, графический дизайн, функции отслеживания, топология, классификаторы, 3d визуализация и многое другое. Особенность библиотеки заключается в том, что каждый ее компонент может использоваться отдельно, не ссылаясь на другие компоненты. Таким образом, в приложении можно использовать только то, что действительно необходимо.

VXL используется по всему миру. Библиотека применяется в сфере обучения и промышленности, некоторые ведущие мировые эксперты в сфере компьютерного зрения пользуются данной библиотекой. Существует документация по VXL с описанием каждого класса и функций. Однако у данной библиотеки существует такой же недостаток, как и у AForge.NET. Разработчику, не знающему английский язык будет весьма проблематично освоить и использовать данный набор библиотек. На рис. 4 приведено иерархическое строение ядра VXL [6].

### **Библиотека LTI**

LTI или LTI-lib – объектно-ориентированная библиотека алгоритмов и структур данных. Она часто применяется при обработке изображений и в сфере компьютерного зрения. LTI-lib была разработана как часть научно-исследовательских проектов в области компьютерного зрения с технологиями робототехники, распознавания объектов, голоса и жестов. Основной целью разработки данной библиотеки является создание объектно-ориентированной библиотеки на языке C++, что во многом упрощало бы использование кода и его обслуживание, но при этом были бы обеспечены быстрые алгоритмы, которые можно было бы использовать в реальных приложениях.

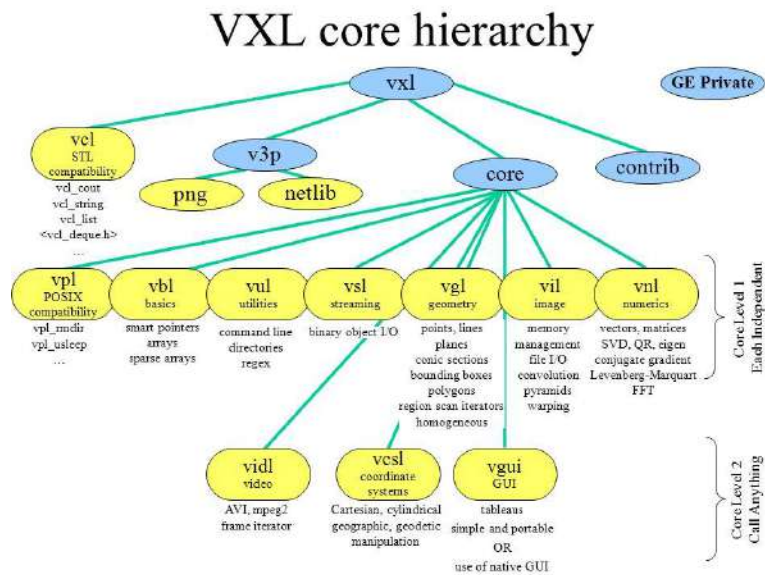


Рисунок 4 – Иерархическое строение ядра VXL

Библиотека была разработана с применением GCC (набор компиляторов, применяемый для разнообразных языков программирования) под Linux и Visual C++ под Windows NT. Многие классы инкапсулируют Windows/Linux функциональность для того, чтобы упростить решение системных или аппаратных задач (например, классы для многопоточности и синхронизации, измерения времени и доступ к последовательному порту) [7].

Библиотека снабжена документацией [8], которая поддерживается разработчиками в актуальном состоянии и может быть найдена на главной странице проекта. Вся информация в ней предоставляется на английском языке, как и в большей части других библиотек. Библиотека LTI, так же как и библиотека OpenCV, является свободно распространяемым программным обеспечением, согласно GNU Lesser General Public License. Существует огромное количество проектов, при разработке которых применялись технологии компьютерного зрения, для которых использовалась именно библиотека LTI.

#### **Система распознавания и классификации изображений транзисторов и выбор библиотеки компьютерного зрения для нее**

Для решения типовых проблем распознавания и классификации изображений транзисторов поставлена задача разработки программной системы с клиент-серверной архитектурой.

Основная часть работы приложения будет выполняться на стороне клиента и заключается в распознавании символической маркировки на транзисторе. Далее, распознанный текст с маркировкой будет передаваться на сервер, на котором находится хранилище (база данных) информации по транзисторам. После поиска соответствий на сервере, информация по транзистору с данной маркировкой возвращается клиенту.

Система должна иметь высокую отказоустойчивость, стабильное подключение к сети Интернет для обмена данными, корректно организованную политику конфиденциальности для доступа к информации на сервере БД. Клиент планируется разрабатывать под платформу Android, как одну из самых распространенных и доступных мобильных систем.

На рис. 5 представлена укрупненная схема работы данной системы.

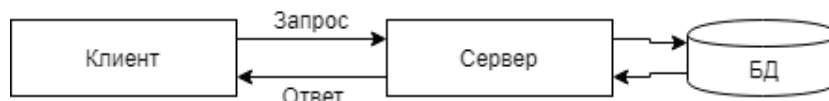


Рисунок 5 – Укрупненная структура системы распознавания и классификации изображений транзисторов

Среди рассмотренных вариантов библиотек компьютерного зрения оптимальным будет использование библиотеки OpenCV, так как она является быстро работающей [1, 2, 5], в ее составе имеются функции, предназначенные не только для распознавания текста, но и для обработки изображения в целом [1, 2, 5], что упрощает структуру будущего приложения и, также имеется реализация библиотеки специально под ОС Android.

## Выводы

При анализе данных о приведенных выше библиотеках компьютерного зрения были получены следующие выводы в пользу использования библиотеки OpenCV для разработки приложения по распознаванию транзисторов.

1. Основным преимуществом библиотеки компьютерного зрения, как и любого другого программного обеспечения, является ее производительность. Как видно из диаграммы на рис. 1, производительность OpenCV превышает производительность аналогов (VXL и LTI-lib) даже без использования дополнительного компонента IPP.

2. Библиотека содержит огромное количество функций для решения разнообразных задач, начиная с обработки изображений, компьютерного зрения и заканчивая обучением машин. Кроме того, библиотека имеет открытый исходный код и лицензию, позволяющую использовать весь функционал для разработки коммерческих продуктов.

3. Преимуществом OpenCV является наличие русскоязычной документации, наличие крупнейшего числа tutorиалов, уроков, научных материалов и книг по применению функционала и методов работы с библиотекой. Нельзя не упомянуть о весьма активном сообществе разработчиков и пользователей библиотеки, которые могут поделиться своим опытом и ответить на интересующие вопросы.

4. Решающим преимуществом в выборе OpenCV стала ее кроссплатформенность и возможность использовать в связке практически с любым языком программирования.

## Литература

1. OpenCV library [Электронный ресурс]. – Режим доступа: <https://opencv.org/>. - Загл. с экрана.
2. OpenCV vs VXL vs LTI: Performance Test - AI Shack [Электронный ресурс]. – Режим доступа: <http://www.aishack.in/tutorials/opencv-vs-vxl-vs-lti-performance-test/>. - Загл. с экрана.
3. AForge.NET :: Framework [Электронный ресурс]. – Режим доступа: <http://www.aforgenet.com/framework/>. - Загл. с экрана.
4. Применение библиотеки AForge.NET и ее расширения Accord.NET Framework при распознавании лиц в режиме реального времени | Статья в журнале «Молодой ученый» [Электронный ресурс]. – Режим доступа: <https://moluch.ru/archive/154/43602/>. - Загл. с экрана.
5. VXL - C++ Libraries for Computer Vision [Электронный ресурс]. – Режим доступа: <http://vxl.sourceforge.net/>. - Загл. с экрана.
6. What is VXL ? Vision Something Libraries A collection of Computer Vision libraries Open Source, grass roots effort, 53 developers –Supported by good community, - ppt download [Электронный ресурс]. – Режим доступа: <https://slideplayer.com/slide/7433547>. - Загл. с экрана.
7. LTI-Lib [Электронный ресурс]. – Режим доступа: <http://ltilib.sourceforge.net/doc/homepage/index.shtml>. - Загл. с экрана.
8. LTI-Lib [Электронный ресурс]. – Режим доступа: <http://ltilib.sourceforge.net/doc/html/index.shtml>. - Загл. с экрана.

*Макогон С.А., Зори С.А. Сравнение популярных библиотек компьютерного зрения для использования в приложении по распознаванию транзисторов. В данной статье произведен обзор и сравнение библиотек OpenCV, AForge.NET, LTI и VXL для использования в приложениях по распознаванию образов и доказана целесообразность использования библиотеки OpenCV для приложений по распознаванию транзисторов.*

**Ключевые слова:** компьютерное зрение, распознавание образов, библиотека, OpenCV, AForge.NET, LTI, VXL, документация

*Makogon S.A., Zori S.A. Comparison of popular computer vision libraries for use in transistor recognition applications. This article reviewed and compared the OpenCV, AForge.NET, LTI and VXL libraries for use in pattern recognition applications and proved the feasibility of using the OpenCV library for transistor recognition applications.*

**Keywords:** computer vision, image recognition, library, OpenCV, AForge.NET, LTI, VXL, documentation