

# Разработка и внедрение АПАК для поиска дефектов изделий микроэлектроники с помощью искусственного интеллекта

## Часть 7. Новая технология поиска дефектов на основе многоракурсной структуры

В ЭК2–7 2025 г. описаны методы и способы настройки изображений для видимого диапазона обнаружения дефектов, рассмотрены методы измерения, классификации и формирования базы данных (БД) дефектов с помощью автоматизированного программно-аппаратного комплекса (АПАК) для поиска дефектов изделий электронной техники (ИЭТ) с сохранением изображения дефекта в БД для дальнейшего применения, описано обнаружение дефектов полупроводниковых пластин в поляризованном свете, рассмотрена проверка качества порошковых материалов и микроструктур поверхностей, описан поиск дефектов микросварки с помощью электромагнитных устройств, приведены результаты исследований в области поиска дефектов путем анализа их ключевых особенностей с использованием современных алгоритмов компьютерного зрения (КЗ) на основе особых точек ORB.

В этой части статьи описывается новая технология поиска дефектов на основе многоракурсной структуры. Результаты экспериментов на изделиях микроэлектроники подчеркивают универсальность и надежность разработанного подхода. Интегрируя многоракурсные прогнозы и совершенствуя методы с обратной связью в реальном времени с КЗ, АПАК способен значительно улучшить обнаружение дефектов. Дальнейшая работа будет сосредоточена на применении представленных результатов исследований к другим типам изделий электронной техники.

**Дарья ДОРМИДОШИНА**,  
заместитель генерального директора,  
АО «ЦКБ «Дейтон»  
**Юрий ЕВСТИФЕЕВ**,  
к.т.н., старший научный сотрудник,  
профессор, АО «ЦКБ «Дейтон»,  
medved55@mail.ru  
**Вячеслав МАЛЫШЕВ**,  
старший инженер, АО «ЦКБ «Дейтон»,  
**Артем НАЗАРЕНКО**,  
старший инженер, АО «ЦКБ «Дейтон»,  
**Юрий РУБЦОВ**,  
генеральный директор,  
главный конструктор,  
АО «ЦКБ «Дейтон»

### Введение

Автоматизированное обнаружение дефектов является важнейшей составляющей контроля качества изделий микроэлектроники. При проведении инспекции особенно сложно идентифицировать дефекты малого размера и дефекты, сливающиеся с фоном поверхности изделий. По этим причинам в статье исследуются и анализируются результаты эффективного многоракурсного подхода машинного обучения (далее МО) для повышения точности обнаружения дефектов путем внедрения и сравнения ранних и поздних методологий слияния. Кроме того, описывается применение новой методологии слияния, которая выравнивает и объединяет многоракурсные признаки с помощью блока преобразования для повышения точности обнаружения дефектов. Для оценки производительности проводились эксперименты на записанном наборе дефектов кристаллов, смонтированных на площадке многовыводных рамок, и сравнивался традиционный одноракурсный контроль с мето-

дами позднего и раннего слияния. Полученные результаты показали, что как ранний, так и поздний методы слияния повышают точность обнаружения по сравнению с одним ракурсом; при этом примененный метод обеспечивает высокие показатели оценки F1. В статье представлены адаптированные метрики точности–полноты, разработанные для обнаружения дефектов на основе сегментации и снимающие ограничения традиционных оценок. Эти метрики точнее отражают обнаружение малоразмерных дефектов. Достиженные результаты подчеркивают преимущества разработанных подходов для проведения визуального контроля качества, обеспечивая надежный и масштабируемый метод обнаружения дефектов изделий микроэлектроники.

Применение компьютерного зрения (КЗ) и МО в системах контроля качества изделий микроэлектроники автоматизировало процессы поиска дефектов, сократило ручной труд, уменьшило ошибки и затраты предприятий. Несмотря на эти достижения, обнаружение малоразмерных дефектов (микродефектов) остается критической про-

блемой. Эти дефекты часто трудно идентифицировать с помощью автоматизированного программно-аппаратного комплекса (АПАК) с одним ракурсом из-за зависимости от углов обзора. Обеспечение тщательного осмотра таких дефектов требует многоакурсного анализа, который снимает изображение изделия с разных точек обзора для повышения точности и надежности их обнаружения.

Одноракурсные АПАК применимы не для всех изделий. Эти комплексы ограничены в своей способности адаптироваться к динамическим требованиям промышленных сред, где такие факторы как геометрически сложные поверхности изделий, нестабильное освещение и микродефекты создают значительные проблемы. Устранение этих ограничений требует инновационного сочетания методов искусственного интеллекта (ИИ) и адаптивных систем визуализации для максимальной точности обнаружения и масштабируемости.

Многоакурсное обнаружение дефектов имитирует процесс человеческого осмотра, исследуя изделия с разных сторон, что позволяет полнее обнаруживать микродефекты. Интегрируя результаты осмотра изделия с разных сторон, многоакурсные АПАК эффективно фиксируют малоразмерные, нерегулярные или скрытые дефекты, которые остаются необнаруженными с определенных точек обзора. Многоакурсное обнаружение обеспечивает целостную оценку поверхности изделия. При этом эффективное объединение информации с нескольких точек обзора требует передовых методов слияния для совместного обоснования данных, обеспечивая точные и последовательные оценки качества изделий с разных точек обзора.

Учитывая проблемы, связанные с нерегулярными и едва заметными дефектами, такие традиционные методы как сегментация фрагментов часто оказываются недостаточными. Следовательно, применяется семантическая сегментация, то есть решается задача КЗ, в которой каждому пикселю изображения присваивается метка класса. Это позволяет лучше выделить на изображении объекты и области, относящиеся к определенным категориям, например к «царапинам», «трещинам», «вмятинам», «выпуклостям» и т.д. В отличие от обычной классификации изображений, которая определяет класс всего изображения, семантическая сегментация обеспечивает более детальное понимание содержимого изображения и делает его основой для разработки технологии обнаружения дефектов.

## Основная часть

В статье представлена новая технология поиска дефектов на основе ИИ, которая использует анализ многоакурсных изображений и передовые методы слияния. Основу этой технологии составляет современное программное обеспечение, эффективно интегрирующее многоакурсную информацию для устранения ограничений подходов с одним видом, что значительно повышает точность обнаружения едва заметных дефектов. Слияния основаны на базовой модели семантической сегментации, благодаря чему повышается их способность локализовать и классифицировать дефекты с большей точностью.

Для обеспечения многоакурсного анализа АПАК включает активную настройку КЗ. Этот комплекс использует роботизированный манипулятор для динамической регулировки точек обзора изделий камерой, захватывая изображения с разных углов для всестороннего визуального охвата, или изделие перемещается перед камерой так, чтобы дефекты были видны с разных точек зрения. Роботизированный манипулятор повышает надежность эффективности АПАК при обнаружении малоразмерных или сложных дефектов. Другим ключевым инструментом технологии является автоматическая аннотация, разработанная для оптимизации трудоемкого процесса аннотации обучающих данных. Используя способность активной КЗ получать многоакурсную информацию, инструмент автоматизирует генерацию аннотаций, проецируя помеченные точки на другие виды, сокращая ручные усилия и обеспечивая согласованность между наборами данных. Такой подход значительно ускоряет подготовку высококачественных наборов данных, которые имеют решающее значение для слияния.

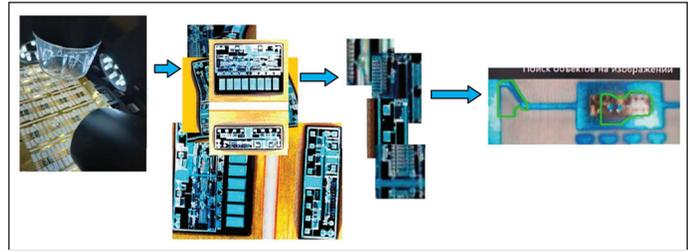


Рис. 1. Ключевые этапы сбора, обработки и анализа данных для обнаружения дефектов

Предлагаемый инструмент ИИ реализован с помощью структурированного АПАК, который объединяет следующие ключевые этапы сбора, обработки и анализа данных (рис. 1):

- 1) сбор данных: АПАК позиционирует камеру для захвата изображений с разных точек обзора, обеспечивая полный охват проверяемого изделия. Или изделие перемещается перед камерой;
- 2) аннотация: захваченные изображения аннотируются для маркировки дефектов с использованием инструмента автоматической аннотации для эффективного создания согласованных высококачественных меток;
- 3) обнаружение дефектов: для обнаружения дефектов используются усовершенствованные модели МО, раннее и позднее слияния в еще большей мере повышают точность обнаружения, используя многоакурсный анализ;
- 4) уточнение прогноза: выходные данные слияния обрабатываются для получения окончательных прогнозов, что обеспечивает точную локализацию и идентификацию дефектов.

Кроме того, вводится новая метрика оценки для устранения ограничений стандартных метрик, обеспечивая более надежную и подробную оценку производительности обнаружения дефектов. Эта новая метрика, специально разработанная для задач обнаружения дефектов, обеспечивает точное сравнение между моделями и подчеркивает возможности применяемого слияния.

## Анализ предшествующих работ

Применение ИИ для обнаружения дефектов претерпело значительные усовершенствования, особенно с появлением МО и сверточных нейронных сетей (СНС). Предыдущие результаты работ продемонстрировали потенциал АПАК на основе ИИ с использованием базовых архитектур нейронных сетей. Хотя эти методы автоматизировали определенные аспекты контроля качества, их одноракурсная природа ограничивала способность обнаруживать малоразмерные или сложные дефекты, которые часто зависят от углов обзора и условий освещения.

Чтобы преодолеть эти ограничения, многообещающим подходом стали методы многоакурсного контроля. Применение СНС в отношении многоакурсных наборов данных в АПАК по примеру [1] показало, что анализ объекта с разных углов позволяет повысить точность обнаружения. Аналогичным образом, наши работы, как и [2], подчеркивали важность захвата разных перспектив для проверки дефектов на поверхностях с высокой отражающей способностью. Однако эти методы часто полагались на фиксированные настройки камеры, что снижало их адаптивность к реальным промышленным средам с разными потребностями в инспекции.

Динамическая настройка точки обзора в еще большей мере продвинула эту область, как показано в [3]. Используя эти идеи, был разработан метод многоакурсного обучения для обнаружения дефектов в полупроводниковых пластинах с помощью поляризационных фильтров [4]. Такой подход улучшил охват проверяемых изделий, но был реализован, в первую очередь, за счет усложнения оборудования, а не оптимизации процессов обнаружения, управляемых ИИ.

Кроме того, для обнаружения дефектов в структурированных средах исследовались роботизированные системы контроля качества.

В [5] описана разработанная автономная роботизированная система для структурного контроля туннелей, интегрирующая датчики и основанные на зрении алгоритмы для обнаружения дефектов. Аналогичным образом, в [6] предложен роботизированный инспектор для оценки туннелей, использующий методы глубокого обучения для выявления поверхностных дефектов. Эти подходы демонстрируют потенциал роботизированного контроля, согласующегося с активной технологией КЗ для обнаружения дефектов.

Интеграция многокурсовых методов с использованием технологий слияния стала еще одной областью интересов. В [7] предложен подход на основе слияния для агрегации информации из нескольких камер, что повысило точность обнаружения дефектов. Однако эта работа, в основном, опиралась на стратегию веса на уровне изображений, которая не полностью использовала взаимодействия между изображениями на уровне признаков. В более ранних работах изучалось слияние на уровне признаков, но эти методы не применялись к сценариям обнаружения дефектов. Наш подход расширяет эту линию исследований, представляя новую архитектуру, которая объединяет стратегии раннего и позднего слияний, адаптированные для обнаружения микродефектов. Этот гибридный подход использует как интеграцию на уровне признаков, так и на уровне решений для обеспечения необходимой производительности, особенно для малоразмерных дефектов.

Более поздние подходы [8], исследовавшие гибридные методы слияния для распознавания многокурсовых объектов, продемонстрировали потенциал объединения стратегий раннего и позднего слияний для повышения надежности. Однако эти методы редко тестировались на сложных дефектах, где мелкие детали имеют решающее значение для точного обнаружения.

Оценочные метрики являются критически важным компонентом исследования для обнаружения дефектов. Несмотря на широкое использование таких стандартных метрик как средняя точность (mean Average Precision, mAP) и пересечение над объединением (Intersection over Union, IoU), они часто не отражают нюансы микродефектов. В [9] представлены метрики, предназначенные для обнаружения дефектов с высоким разрешением и обеспечивающие точность на уровне пикселей. Эти достижения подчеркивают необходимость в метриках, ориентированных на конкретные задачи, которые наши исследования решают посредством разработки новой оценочной метрики для обнаружения малоразмерных дефектов.

Основываясь на этих достижениях, настоящие исследования представляют комплексную структуру обнаружения дефектов на основе ИИ, которая объединяет многокурсовую инспекцию и новую оценочную метрику, адаптированную под малоразмерные дефекты. В отличие от предыдущих работ, которые подчеркивали статические

многокурсовые настройки или общие метрики, результаты настоящего исследования объединяют передовые методы ИИ с адаптивными системами визуализации для повышения точности и масштабируемости. Благодаря такому объединению устраняются критические пробелы в многокурсном обнаружении дефектов, что значительно улучшает выявление сложных дефектов и контроль над качеством изделий микроэлектроники.

В то время как многокурсовые и роботизированные методики контроля качества улучшают обнаружение дефектов, промышленные условия эксплуатации создают дополнительные проблемы. Изменение освещения, отражательная способность поверхности могут потребовать дополнительных точек обзора для точной идентификации дефектов. Кроме того, баланс между точностью обнаружения и эффективностью обработки в реальном времени имеет решающее значение для развертывания в промышленных условиях. Ограничения движения роботизированной части АПАК дополнительно влияют на выбор точки обзора, что требует оптимизации для полного покрытия площади проверки. Такой подход объединяет методы многокурсового анализа и слияния для решения этих задач, обеспечивая надежность и эффективность обнаружения дефектов.

## Методология исследования

### Обзор обнаружения дефектов с помощью ИИ

Процесс обнаружения дефектов с помощью ИИ начинается с надежного сбора данных с использованием активного КЗ. АПАК оснащен манипулятором, на которой установлена камера (рис. 2).

Инспектируемое изделие устанавливается на стол, позволяя АПАК захватывать многокурсовые изображения с разных углов. Для оптимального освещения, выбираемого с помощью соответствующих настроек, в камеру вмонтированы источники света.

Первым шагом является сбор данных, когда КЗ захватывает изображения с разных углов, используя автоматическую оптимизацию точек обзора. Этот процесс является ключевым для осмотра всей поверхности изделия с разных точек.

### Оптимизация точек обзора

В современном производстве тенденции к массовой настройке и высоким требованиям к качеству создают уникальные проблемы. Управляемый человеком АПАК обеспечивает гибкость, в отличие от последовательности и повторяемости. АПАК, использующий оснащенные камерами манипуляторы, предоставляет многообещающее решение. АПАК должен решать такие ключевые проблемы как оптимальный выбор точки обзора и эффективное планирование пути для всестороннего охвата проверяемых поверхностей.

Для этого в представленных исследованиях используются методы [10], которые обеспечивают планирование пути покрытия с целевой выборкой точек обзора для роботизированного контроля поверхностей изделия. Такой подход оптимизирует выбор точки обзора и планирование пути, чтобы минимизировать время цикла, предоставляя при этом полный охват для инспекции. Основываясь на этих концепциях, мы адаптируем упомянутые методы к нашей настройке КЗ для захвата многокурсовых изображений с целью полной проверки поверхности, особенно в случае сложных дефектов.

### Автоматическая аннотация

Одной из целей настоящих исследований было значительное сокращение ручных усилий, необходимых для создания больших аннотированных наборов данных для обнаружения дефектов. Для создания аннотаций нескольких изображений дефектного изделия были оптимизированы процессы, и существенно поднята эффективность. Это стремление обусловлено признанием того обстоятельства, что модели СНС МО требуют обширных наборов данных для достижения высокой точности. Однако аннотация таких данных является трудоемкой и часто требует экспертных знаний, быстро становясь узким местом в рабочем процессе.

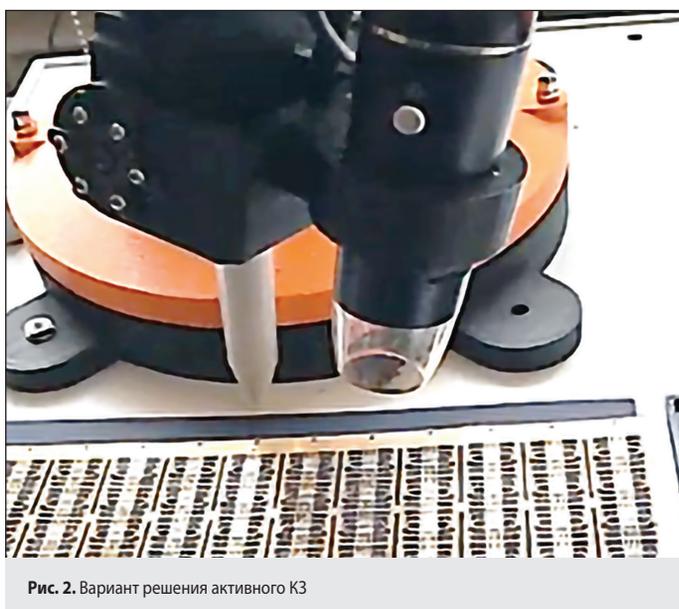


Рис. 2. Вариант решения активного КЗ

Настоящий подход использует камеру для захвата изображений с разных точек обзора изделия. Используя кинематику манипулятора процесса калибровки, точно вычисляется позиция объектива камеры относительно изделия. Инструмент автоматической аннотации позволяет выбирать и аннотировать одно изображение. По завершении из этой аннотации генерируется маска сегментации, которая затем применяется для каждого изображения в наборе данных.

Результаты такого подхода позволяют создавать наборы данных со значительно меньшими усилиями, чем при ручном аннотировании. Одно ручное аннотирование может быть усилено для создания аннотаций для двадцати изображений, что значительно сокращает время аннотирования. Такая эффективность не только облегчает быстрый сбор новых наборов данных, но и повышает способность изучать и обнаруживать новые дефекты поверхности, что является значительным достижением в области КЗ.

## Обработка данных

### Предварительная обработка

В КЗ АПАК предварительная обработка преобразует необработанные данные в специальный формат для лучшего обучения модели. Процесс начинается с удаления фона, выделяя объекты, на которых ИИ может сосредоточиться. ИИ также упрощает данные, удаляя ненужные детали. Затем изображение обрезается, чтобы ИИ сосредоточился на самых важных частях или областях. При этом ИИ уделяет внимание ключевым местам, эффективнее используя свои ресурсы. Затем размер изображений делается стандартным, что помогает ИИ обрабатывать их быстрее. Это не значит, что теряются важные детали; речь идет о поиске правильного баланса для эффективности и ясности. Шаги, начиная с удаления фона и заканчивая изменением размера, направлены на максимальное упрощение данных. Таким образом, АПАК может учиться на лучших возможных примерах, что делает его «умнее и точнее» в понимании и интерпретации визуальной информации.

### Дополнение

После предварительной обработки АПАК переходит к дополнению данных – важному шагу для повышения надежности и адаптивности процесса. Используется искусственное расширение полученного набора данных и повышение устойчивости полученной модели к разным визуальным сценариям. Одним из основных дополнений является горизонтальное отражение, которое зеркально отображает изображения относительно вертикальной оси и имитирует изменчивость ориентации объекта, с которой может столкнуться АПАК. Кроме того, вводятся повороты до определенных градусов на основе данных, чтобы имитировать изменения наклона и угла в практических наблюдениях, благодаря чему повышается способность модели распознавать объекты с разных углов зрения. Корректировки цвета, яркости и контрастности также применяются случайным образом. Эти изменения помогают модели научиться распознавать объекты в разных условиях освещения и цветовых вариациях, обеспечивая постоянную производительность независимо от визуальной среды. В дополнение к этим преобразованиям на этапе предварительной обработки используется фактор случайности. Например, намеренно вносятся небольшие ошибки при удалении фона. Это может показаться нелогичным, но учит модели адекватно воспринимать недостатки и недостающую фокусировку изображения изделия.

Такие шаги предварительной обработки и дополнения гарантируют, что АПАК работает на чистых, согласованных и разнообразных наборах данных, что обеспечивает их надежность и точность.

## Применяемые технологии нейронных сетей

### Семантическая сегментация

Решение использовать семантическую сегментацию проистекает из ее способности обрабатывать микродефекты, которые имеют

неправильную форму или значительно различаются по размеру. Семантическая сегментация отлично обнаруживает и локализует неправильные дефекты, классифицируя каждый пиксель, что делает ее особенно эффективной для сложных и малозаметных дефектов.

Семантическая сегментация – задача КЗ, в которой каждому пикселю изображения присваивается метка класса. Это позволяет выделить на изображении объекты и области, относящиеся к определенным категориям, например к «царапинам», «трещинам», «вмятинам», «выпуклостям» и т.д. В отличие от обычной классификации изображений, которая определяет класс всего изображения, семантическая сегментация делает это на уровне отдельных пикселей, обеспечивая более детальное понимание содержимого изображения.

Модели вроде Fully Convolutional Networks (FCN), U-Net и DeepLab были специально разработаны для того, чтобы обрабатывать каждую деталь изображения на уровне пикселей. Было исследовано, как эти модели работают, что отличает их друг от друга и почему они стали ключевыми инструментами для КЗ.

Модели сегментации:

- 1) UNet – широко используемая сверточная сеть, которая разработана специально для сегментации изображений. Она принята для решения задач семантической сегментации, включающих сегментацию дефектов. UNet следует симметричной структуре кодер–декодер, в которой кодер захватывает иерархические признаки с помощью последовательной понижающей дискретизации, в то время как декодер постепенно повышает дискретизацию карт признаков для восстановления пространственного разрешения. Ключевым аспектом UNet являются ее пропускные соединения, которые напрямую передают низкоуровневые признаки от кодера к декодеру, сохраняя мелкозернистую пространственную информацию и повышая точность сегментации;
- 2) FCN улучшает представление признаков, используя подход сверху вниз и латеральные связи. В отличие от традиционных архитектур, которые полагаются исключительно на самые глубокие карты признаков, FCN строит многомасштабную пирамиду признаков, интегрирующую не только высокоуровневую семантическую информацию, но и мелкозернистые пространственные детали. Такая структура позволяет лучше сегментировать объекты в разных масштабах, что делает ее особенно эффективной для анализа сложных и многомасштабных изображений;
- 3) DeepLabv3+ расширяет семейство DeepLab, включая расширенные свертки и улучшенный модуль декодера. Расширенные свертки обеспечивают большее рецептивное поле без увеличения вычислительной сложности, позволяя сети эффективно захватывать контекстную информацию. Добавление явного декодера позволяет также улучшить результаты сегментации, восстанавливая пространственные детали, потерянные при понижении частоты дискретизации. DeepLabv3+ известна своей надежностью при сегментации объектов со сложными границами и разными размерами.

Такое сравнение служит эталоном в наших исследованиях для поиска лучшей базовой сети и определения методов слияния.

### Раннее и позднее слияние

В области КЗ, особенно в приложениях по обнаружению дефектов, требующих многовидового анализа, интеграция информации с нескольких точек зрения имеет решающее значение для повышения точности и надежности. Применяются две основные методологии интеграции многоакурсной информации – позднее и раннее слияние. Оба подхода имеют разные преимущества и операционные механизмы для объединения данных из нескольких источников для улучшения общего процесса принятия решений СНС.

Позднее слияние относится к методу, при котором отдельные представления данных обрабатываются независимо через параллельные процессы, а их выходные данные объединяются в конце обработки. Такой подход позволяет выполнять специализированную обработку каждого представления, используя уникальные особенности или характеристики перед принятием решения. Основное преимущество

позднего слияния заключается в его простоте и гибкости, которую оно предлагает при обработке и интеграции разнородных источников данных. Архитектура позднего слияния показана на рис. 3, который демонстрирует отдельную обработку для каждого представления кристаллов многокристального модуля и их слияние на уровне принятия решения.

Раннее слияние подразумевает интеграцию данных из разных представлений на ранней стадии, до или во время начальных слоев обработки СНС. Снабжая СНС составным представлением, объединяющим несколько точек осмотра изделия, раннее слияние позволяет модели использовать взаимосвязи между разными представлениями данных с самого начала. Предполагается, что этот подход способствует более целостному пониманию данных, что потенциально приводит к более точным и надежным обнаружениям, позволяя СНС анализировать данные более унифицированным образом. На рис. 4 представлен обзор архитектуры раннего слияния, где данные из нескольких представлений кристаллов многокристального модуля объединяются на уровне признаков перед последующей обработкой для обнаружения дефектов.

Настоящее исследование углубляется в сравнительную эффективность этих методологий слияния в области многоакурсного обнаружения дефектов. Предполагается, что способность раннего слияния к одновременному рассмотрению нескольких представлений может дать более надежные результаты обнаружения. Чтобы реализовать раннее слияние, представляется новая архитектура, специально применяемая для выравнивания глубоких признаков из разных представлений.

Такая архитектура позволяет обрабатывать входные данные несколькими представлениями. Она состоит из N (где N – количество представлений) кодировщиков и одного декодера. В традиционной архитектуре признаки блока кодировщика объединяются с признаками блока декодера. В рассматриваемом случае это не так просто, поскольку обрабатываются признаки с несколькими представлениями, которые не выровнены. Чтобы решить эту проблему, включается новый блок слияния (Т-блок), который преобразует и объединяет многовидовые признаки N кодеров, после чего они объединяются с признаками декодера (рис. 4). Этот Т-блок состоит из модуля преобразования перспективы, выполняющего перспективное преобразование (гомографию) признаков, выравнивая их с опорным видом. В данном случае признаки выравниваются с одним из многовидовых входов. Учитывая наличие полностью откалиброванной установки, можно вычислить требуемые параметры для Т-блоков, поскольку имеется информация о местоположениях камеры относительно изделия.

В отличие от раннего слияния, подход позднего слияния использует отдельные каналы обработки для каждого вида, интегрируя результаты обнаружения только на последних этапах с помощью механизма слияния. Сравнивая обе методологии, достигается понимание того, как наиболее эффективно можно использовать многоакурсную ин-

формацию для обнаружения дефектов, а также для разработки более сложных и надежных АПАК.

**Метрики оценки**

Одним из важных аспектов оценки методов МО для обнаружения дефектов изделий микроэлектроники является тщательный выбор метрик, которые отражают производительность модели, будучи при этом релевантными для конкретного варианта использования. Для обнаружения малоразмерных дефектов широко используются такие метрики обнаружения как точность, полнота и их объединенная метрика – оценка F1. Они предпочтительнее более сложных метрик, например mAP, потому что проще для понимания и напрямую связаны со стандартами качества.

Для определения метрик вводится двумерное пространство вариантов сопоставления данных: Пл – истинных положительных дефектов; Лп – ложных положительных дефектов; Ло – ложных отрицательных и От – истинных отрицательных дефектов. При этом учитывается тот факт, что не все дефекты являются браком. Такие дефекты должны быть учтены; при этом у изделий снижается качество.

Точность позволяет измерить, сколько из обнаруженных дефектов являются фактическими, а полнота – оценить, сколько из фактических дефектов были правильно обнаружены моделью. Точность математически выражается следующим образом (1):

$$T = \frac{Пл}{Пл + Лп + Ло + От} \tag{1}$$

Полнота (2):

$$П = \frac{Пл}{Пл + Ло} \tag{2}$$

Оценка F1 представляет собой среднее гармоническое значение точности и полноты, уравновешивающее компромисс между ними (3):

$$F1 = 2 \cdot \frac{T \cdot П}{T + П} \tag{3}$$

Хотя эти метрики полезны, они не лишены ограничений, особенно в контексте обнаружения малоразмерных дефектов. Одной из ключевых проблем является неоднозначность, возникающая из-за маркировки истинных данных и строгих правил, используемых для сопоставления прогнозов с примерами истинных данных. В методах, основанных на обнаружении, метрика IoU часто используется для измерения перекрытия между предсказанными и ограничивающими рамками истинных данных. Однако такой подход создает проблемы, когда сама метка истинности данных неоднозначна. Например, длинная царапина может быть помечена как один дефект или может быть разделена на две более мелкие царапины в зависимости от того,

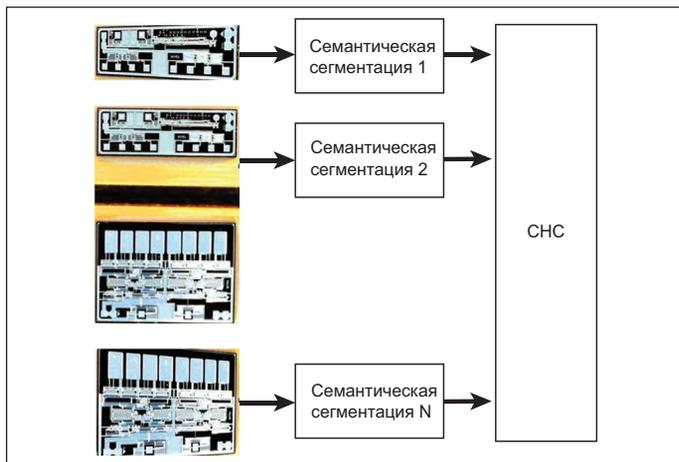


Рис. 3. Архитектура позднего слияния, демонстрирующая отдельную обработку для каждого представления кристаллов многокристального модуля и их слияние на уровне принятия решения

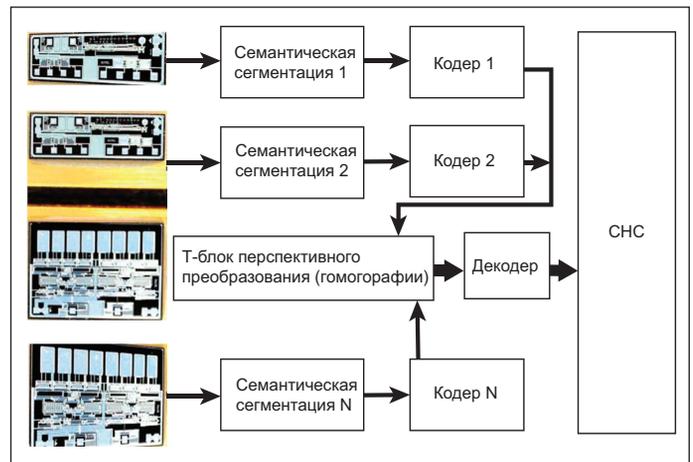


Рис. 4. Архитектура раннего слияния, где данные из нескольких представлений кристаллов многокристального модуля объединяются на уровне признаков перед последующей обработкой для обнаружения дефектов

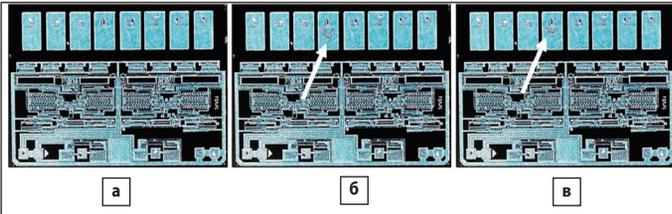


Рис. 5. а) ограничения метрики IoU при обнаружении микродефектов: б) тонкие 0,30 или в) короткие 0,60 прогнозы указывают на низкий показатель IoU, несмотря на правильное обнаружение микродефектов кристалла

как аннотатор (специалист, занимающийся разметкой данных для обучения моделей МО) решит определить царапину. Аналогично, если модель обнаруживает царапину как два более мелких дефекта, традиционные метрики точности и полноты будут занижать результат.

Чтобы преодолеть эту проблему, в наших исследованиях используются методы сегментации. Сегментация, по своей сути, позволяет избегать некоторых двусмысленностей, связанных с маркировкой ограничивающих рамок, поскольку работает на уровне пикселей. Однако возникает вторая проблема – поддержка традиционных метрик точности–полноты в контексте методов сегментации, которые используют маски истинности на основе пикселей, а не дискретные экземпляры.

Типичная метрика, используемая для задач сегментации, – IoU, которая вычисляет пересечение над объединением между предсказанными и истинной масками. Это эффективная метрика может быть слишком строгой для рассматриваемого варианта использования. IoU снижает показатель из-за расхождения в форме или толщине предсказанной маски по сравнению с истинной маской. В случае обнаружения схожих микродефектов небольшие различия в их размерах могут быть несущественными, но, тем не менее, приведут к низкому показателю IoU. Проблема в том, что размеры дефектов менее важны факта обнаружения самого дефекта. Например, немного более толстая или более тонкая предсказанная маска по сравнению с маской истинности все еще может правильно идентифицировать дефект, но показатель IoU снизится, как видно из рис. 5.

Такая строгость метрики IoU для сегментации проблематична в тех случаях, когда обнаружены микродефекты, в которых точные размеры оказываются не столь критичными. Чтобы решить эту проблему, предлагается модификация традиционных метрик точности и полноты, адаптированная к контексту сегментации. Эти метрики сохраняют суть исходных вычислений точности и полноты, но лучше подходят для пиксельной природы сегментации.

Для расчета точности бинарная маска предсказания преобразуется в отдельные компоненты (с помощью алгоритма связанных компонентов), и для каждого компонента оценивается их перекрытие с бинарным истинным значением. Перекрытие вычисляется как отношение пересечения между компонентом и истинным значением к общей площади компонента. Если перекрытие превышает заранее установленный порог, экземпляр считается правильно предсказанным. Затем точность вычисляется как отношение правильно предсказанных компонентов к общему количеству компонентов (4):

$$T = \frac{\text{правильно предсказанные компоненты}}{\text{общее число компонентов}} \quad (4)$$

Далее бинарная маска истинных данных преобразуется в отдельные компоненты, и оценивается их перекрытие с маской бинарного прогнозирования. Перекрытие вычисляется как отношение пересечения между маской бинарного прогнозирования и компонентом истинных данных к общей площади компонента истинных данных. Если перекрытие превышает установленный порог, дефект считается правильно обнаруженным. Полнота вычисляется как отношение правильно обнаруженных компонентов истинных данных к общему количеству компонентов истинных данных (5):

$$P = \frac{\text{правильно обнаруженные компоненты истинных данных}}{\text{общее количество компонентов истинных данных}} \quad (5)$$

Оценка F1 выводится из точности и полноты аналогично (3), но с учетом (4–5); в результате получаем (6):

$$F1 = 2 \cdot \frac{T \cdot P}{T + P} \quad (6)$$

Эта модификация выражений (от 1–3 к 4–6) гарантирует, что метрики оценки сосредоточены на значимой точности и полноте обнаружения; при этом избегается ненужное снижение показателя, вызванное небольшими расхождениями в площади дефектов. Адаптируя традиционные метрики точности–полноты для сегментации, можно эффективно оценивать модели в задачах обнаружения дефектов, которые не чрезмерно чувствительны к вариациям на уровне пикселей.

Эти метрики, адаптированные к задаче сегментации, являются более надежной оценкой моделей обнаружения микродефектов с учетом практических требований независимо от незначительных вариаций площадей или форм дефектов.

## Результаты

В этом разделе оценивается производительность АПАК на наборе данных кристаллов микросхем. Набор данных состоит из шести разделенных кристаллов, закрепленных на площадках матрицы многовыводных рамок, каждый из которых снят 64 изображениями – 32 с правой стороны большей ширины кристаллов и 32 – с левой стороны.

Набор данных создает типичные проблемы при контроле качества. В частности, они касаются обнаружения едва заметных дефектов. Основной целью является проверка эффективности результатов технологий многокурсного раннего и позднего слияний по сравнению с однокурсной семантической сегментацией для обнаружения дефектов.

## Проведенные эксперименты

Все инструменты, охватывающие модели семантической сегментации и слияния, обучаются до конвергенции с использованием функции потерь бинарной кросс-энтропии (функция весов модели для минимизации ошибок предсказания). Скорость обучения установлена на уровне  $5 \cdot 10^{-4}$  (гиперпараметр, который определяет размер шага при обновлении весов модели в процессе обучения). Каждый набор данных разделен на 80% обучающих и 20% тестовых подмножеств для комплексной оценки. Производительность моделей оценивается с использованием точности, полноты и оценки F1, которые адаптированы, как описано выше, к задачам обнаружения дефектов.

Время, необходимое для захвата многокурсных изображений, зависит от таких факторов как количество точек обзора, скорость движения камеры и эффективность обработки. КЗ оптимизировано с предварительно откалиброванными точками обзора для уменьшения повторного позиционирования и повышения скорости съема изображения. Для АПАК сканирование занимает  $0,5 \text{ мм}^2/\text{с}$ . Хотя получение многокурсного изображения вносит дополнительные расходы по сравнению с однокурсным подходом, выигрыш в точности обнаружения дефектов оправдывает этот компромисс. Кроме того, ранняя сеть слияния разработана для работы с минимальным количеством точек обзора, что обеспечивает баланс между точностью и временем съема изображения. Общее время обработки данных существенно зависит от количества точек обзора и вычислительной инфраструктуры.

Сначала сравнивается оценка разных моделей сегментации в наборе данных. Результаты показаны в табл. 1. Модель DeeplabV3+ лучше всего справляется с задачей обработки. За ней следуют модели UNet и FCN. Хотя модель DeeplabV3+ работает лучше всего, продолжен эксперимент по слиянию с UNet, поскольку UNet обеспечивает хороший баланс между сложностью и точностью.

В табл. 2 сравнивается производительность семантической сегментации с одним представлением (UNet), с производительностями моделей раннего и позднего слияний.

**Таблица 1.** Сравнение моделей одноракурсной обработки

Модель сегментации	Точность	Полнота	Оценка F1
UNet	<b>0,928</b>	0,875	0,901
FCN	0,85	0,85	0,85
DeepLabV3+	0,92	<b>0,91</b>	<b>0,91</b>

**Таблица 2.** Сравнение производительностей разных моделей

Модель сегментации	Результат		
	точность	полнота	оценка F1
UNet	0,928	0,875	0,901
раннее слияние	<b>0,944</b>	0,940	<b>0,942</b>
позднее слияние	0,888	<b>0,990</b>	0,936

В табл. 1–2 полужирным шрифтом выделены лучшие результаты. Эксперименты показали следующее:

- 1) UNet: несмотря на то, что эта модель обеспечивает высокую точность 0,928, ее полнота ограничена значением 0,875, что приводит к оценке  $F1 = 0,901$ . Таким образом, эта модель полностью не охватывает все дефекты из-за ограниченной перспективы;
- 2) раннее слияние: обеспечивается наивысший общий показатель  $F1$  (0,942) с хорошим балансом между точностью (0,944) и полнотой (0,940). Таким образом, раннее слияние особенно эффективно в интеграции многокурсовых данных на уровне признаков и обеспечивает надежные результаты обнаружения;
- 3) позднее слияние: полнота значительно улучшилась (0,990) по сравнению с UNet, что указывает на способность эффективно агрегировать прогнозы UNet. Однако точность немного снижается (0,888), что приводит к оценке  $F1 = 0,936$ .

Ниже приведены основные результаты исследований:

- 1) полный съем изображения: надежное обнаружение дефектов требует полного представления изделия с осмотром его с нескольких точек. Как ранний, так и поздний подходы к слиянию значительно превосходят метод с однокурсным изображением;
- 2) сравнение моделей слияния: позднее слияние превосходит по полноте, что делает его подходящим для условий, где критически важно максимальное покрытие дефектов. Раннее слияние достигает наивысшей точности и оценки  $F1$ , что указывает на его способность обеспечить сбалансированную и надежную производительность обнаружения дефектов;
- 3) проблемы раннего слияния: хотя раннее слияние достигает приемлемых результатов, оно является более сложным в вычислительном отношении из-за необходимости точного выравнивания признаков на нескольких изображениях.

**Выводы**

Основные выводы проведенных исследований:

- 1) модели раннего и позднего слияния интегрируют многокурсовую информацию для значительного повышения точности обнаружения, особенно для малоразмерных или сложных дефектов;
- 2) представлена новая метрика оценки, адаптированная для задач обнаружения дефектов, устраняющая ограничения стандартных метрик и обеспечивающая более надежные оценки производительности;
- 3) показаны процессы ИИ, которые интегрируют многокурсовый анализ изображений, аннотацию и уточненные прогнозы для обнаружения малоразмерных или сложных дефектов.

Эти исследования демонстрируют потенциал интеграции методов многокурсового МО для повышения точности обнаружения дефектов. Использование методов раннего и позднего слияний показало, что многокурсовый анализ превосходит традиционные подходы с одним видом при обнаружении малоразмерных дефектов. Особенно эффек-

тивные методы раннего слияния обеспечивают наивысший показатель  $F1$  за счет выравнивания признаков из нескольких изображений, что приводит к более надежной и комплексной сегментации дефектов.

В дополнение к передовым методам слияния разработаны методы КЗ, которые автоматически позволяют настраивать точку обзора для оптимального обнаружения дефектов. Эта настройка в сочетании с аннотированием значительно сокращает ручной труд, обеспечивая быстрое создание высококачественных наборов данных для обучения. Эти инновации оптимизируют процесс МО, делая его более эффективным и масштабируемым для поиска дефектов.

Разработана адаптированная метрика точности–полноты специально для задач сегментации, позволяющая устранить общие ограничения традиционных метрик обнаружения, например IoU, особенно в контексте микродефектов. Эта коррекция гарантирует, что небольшие расхождения в границах или площади дефекта не окажут отрицательного влияния на оценку производительности, обеспечив более надежную оценку возможностей АПАК.

Результаты экспериментов на изделиях микроэлектроники, а также запланированное расширение на другие типы изделий электронной техники подчеркивают универсальность и надежность разработанного подхода. Интегрируя многокурсовые прогнозы и совершенствуя методы с обратной связью в реальном времени с КЗ, АПАК способен значительно улучшить обнаружение дефектов. Дальнейшая работа будет сосредоточена на применении представленных результатов исследований к изделиям электронной техники других типов.

Последующие исследования будут направлены на уточнение оптимизации динамической точки обзора и дальнейшее снижение воздействия шума окружающей среды, освещения и ограничений оборудования для гарантии масштабируемости и эффективности АПАК с целью эксплуатации на предприятиях радиоэлектронной промышленности.

**Литература**

1. Steitz J.-M. O., Saeedan F., Roth S. Multi-view X-ray R-CNN. Proceedings of the German Conference on Pattern Recognition (GCPR). Stuttgart. Germany. 4 October. 2018.
2. Li H., Huang J., Gao M., Yang L., Bao Y. Multi-View Information Fusion Fault Diagnosis Method Based on Attention Mechanism and Convolutional Neural Network. Appl. Sci. 2022.
3. Wu H., Zheng K., Sfarra S., Liu Y., Yao Y. Multiview Learning for Subsurface Defect Detection in Composite Products: A Challenge on Thermographic Data Analysis. IEEE Trans. Ind. Inform. 2020.
4. Дормидошина Д. А., Евстифеев Ю. А., Рубцов Ю. В. Разработка и внедрение АПАК для поиска дефектов изделий микроэлектроники с помощью искусственного интеллекта. Часть 3. Обнаружение дефектов полупроводниковых пластин в поляризованном свете. Электронные компоненты. №4. 2025.

ООО  
СМР



ИНТЕРНЕТ-МАГАЗИН  
**www.SMD.ru**

электронные  
**для поверхностного  
 монтажа**

**НОВОЕ В ПРОГРАММЕ ПОСТАВОК**

- Керамические конденсаторы до 100 мкф
- Синфазные дроссели на ток 10 А

РЕКЛАМА




Москва, Ленинградский пр., 80 к. 32; e-mail: sale@smd.ru  
 Тел.: (499) 158-7396, (495) 940-6244, (499) 943-8780

5. Loupos K., Stentoumis C., Pitsikalis V., Botsialas A., Georgopoulos A., Georgoulas G., Gasteratos A., Amditis A. Autonomous Robotic System for Tunnel Structural Inspection and Assessment. *Int. J. Intell. Robot. Appl.* 2017.
6. Stentoumis C., Loupos K., Botsialas A., Georgopoulos A., Gasteratos A., Amditis A. Autonomous Robotic Inspection in Tunnels. *ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci.* 2016.
7. Jiang X., Guo K., Lu Y., Yan F., Liu H., Cao J., Xu M., Tao D. CINFormer: Transformer Network with Multi-Stage CNN Feature Injection for Surface Defect Segmentation. *arXiv. Computer Vision and Pattern Recognition.* 2023.
8. Chen Z., Feng X., Liu L., Zhang H., Wang J. Surface Defect Detection of Industrial Components Based on Vision. *Sci. Rep.* 2023.
9. Islam M. R., Zamil M. Z. H., Rayed M. E., Kabir M. M., Mridha M. F., Nishimura S., Shin J. Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes. *IEEE Access.* 2024.
10. Glorieux E., Franciosa P., Ceglarek D. Coverage Path Planning with Targeted Viewpoint Sampling for Robotic Free-Form Surface Inspection. *Robot. Comput.-Integr. Manuf.* 2019.