



From Human Subjectivity to Objective Output: Leveraging Robotics for Accurate Cosmetic Grading of Used Smartphones

A guide to implementing device processing sophistication

From Human Subjectivity to Objective Output: Leveraging Robotics for Accurate Cosmetic Grading of Used Smartphones

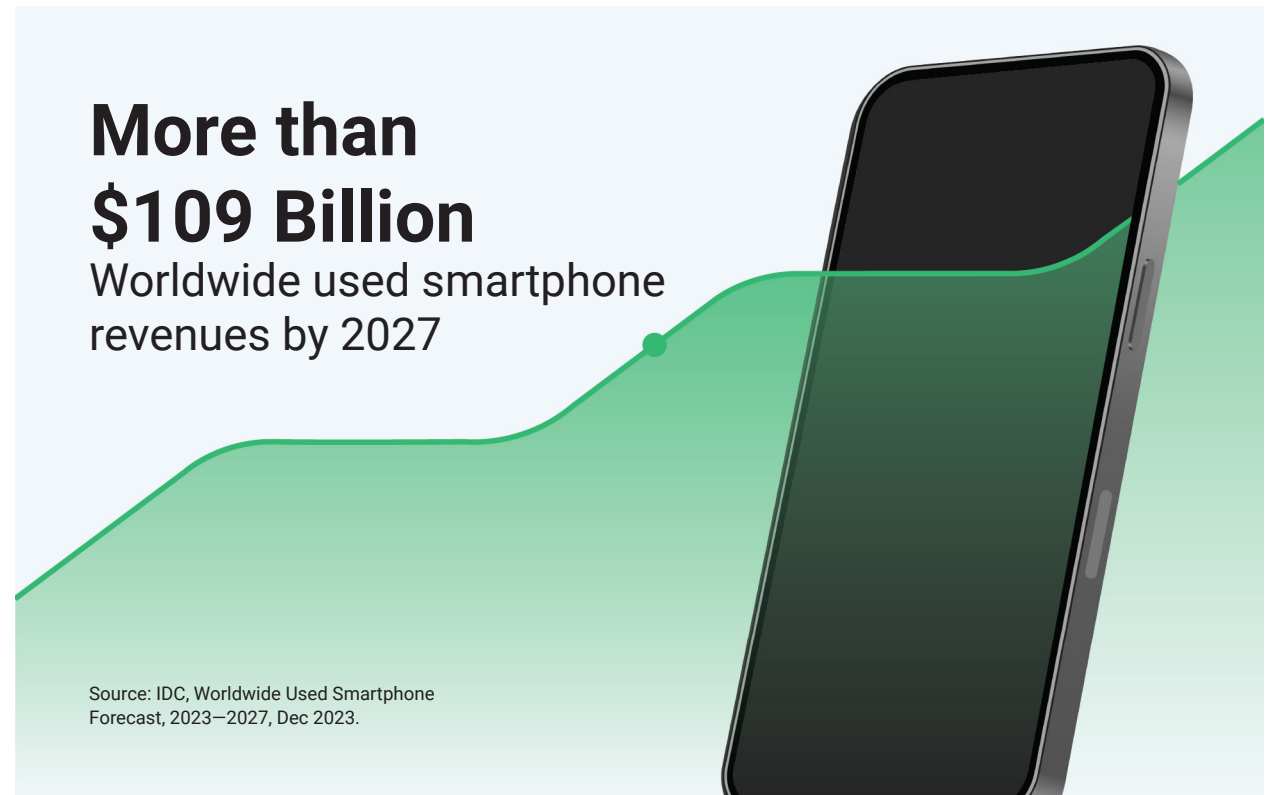
A guide to implementing device processing sophistication



Executive Overview

According to IDC, the global market for used smartphones is expanding rapidly. It is projected to exceed 430 million units and reach a value of \$109.7 billion by 2027. Numerous companies are involved in assessing, refurbishing, and reselling pre-owned devices.

Becoming a dominant player in the secondary device market is a compelling proposition, proving that circularity can be good for business.



Key Terms

A **defect** is any physical or functional imperfection affecting a device's performance, appearance, or usability. Defects are classified into various categories, such as:

- **Cosmetic:** scratches, dents, discoloration, etc.
- **Functional:** touchscreen malfunctions, battery degradation, faulty buttons, etc.
- **Structural:** warped frame, cracked screens, etc.

Defect detection is the **process** of identifying and assessing these defects, often using automated or robotic systems.

Functionality testing assesses the overall performance of essential device functions such as calling, texting, network connectivity, screen responsiveness, camera quality, and battery life.

Cosmetic grading assesses a device's cosmetic condition, including scratches, cracks, and dents. Devices are typically assigned a grade on a scale from A to D, with Grade A being the best condition and Grade D being the worst. The fewer defects detected on a device, the higher the assigned grade.

Smartphone assessments typically involve two key aspects: functionality testing and cosmetic grading.

While functional testing is relatively straightforward, cosmetic grading involves both defect detection and grade assignment. The number of defects detected significantly influences the device's assigned grade and resale value.

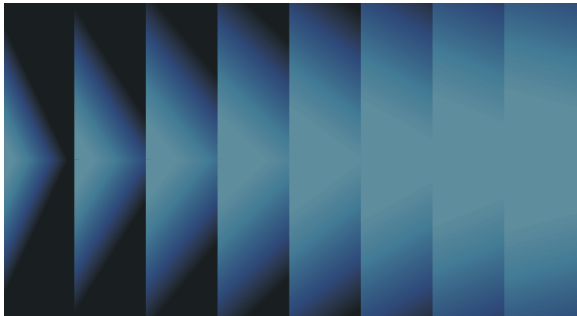
Accurate cosmetic grading ensures fair market value pricing, reduces operational costs, and enhances consumer trust in the secondary smartphone market by ensuring the device a buyer receives is of expected quality. Conversely, poor grading practices can harm profitability, increase customer dissatisfaction, and erode trust within the industry.

A significant challenge in defect detection—and, therefore, cosmetic grading—is the inherent subjectivity of human evaluations. When two human evaluators assess the same device, there are often discrepancies in how many defects each human detects, leading to inconsistent grade assignments.

Minimizing the variance in defect detection is crucial for achieving consistent, reliable standards in device grading. The most effective and efficient way to achieve these consistent results is by using automation to complete defect detection while leveraging the skills of human grading experts to inform the assessment algorithm, monitor the robotic processing automation, and focus on device exceptions.

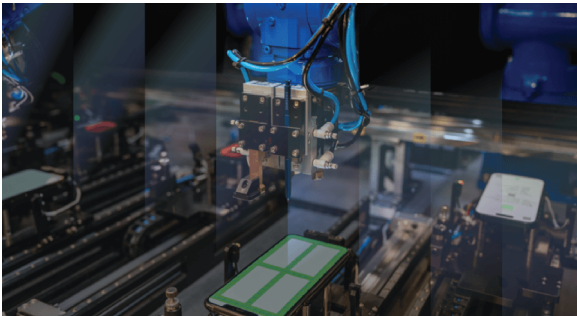


Table of Contents



01
Introduction

PAGE 04



02
The Role of Robotics in the Cosmetic
Grading Process

PAGE 05



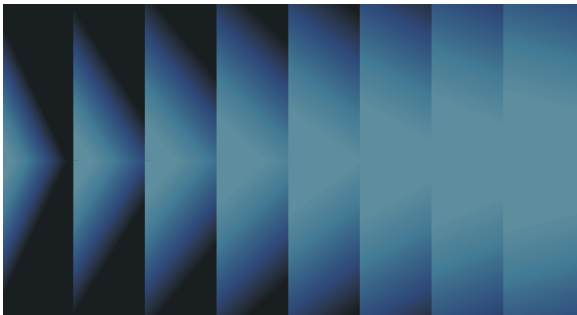
03
The Apkudo Cosmetic Grading
Alignment Method

PAGE 06



04
Keys to Success

PAGE 10



05
Conclusion

PAGE 12



Introduction

In several global regions, the number of pre-owned smartphones sold exceeds the number of new devices sold, with some markets even seeing more revenue from used phones than new ones. Large telecom companies and third-party refurbishment vendors are adapting their processes to meet this increased demand for pre-owned devices.

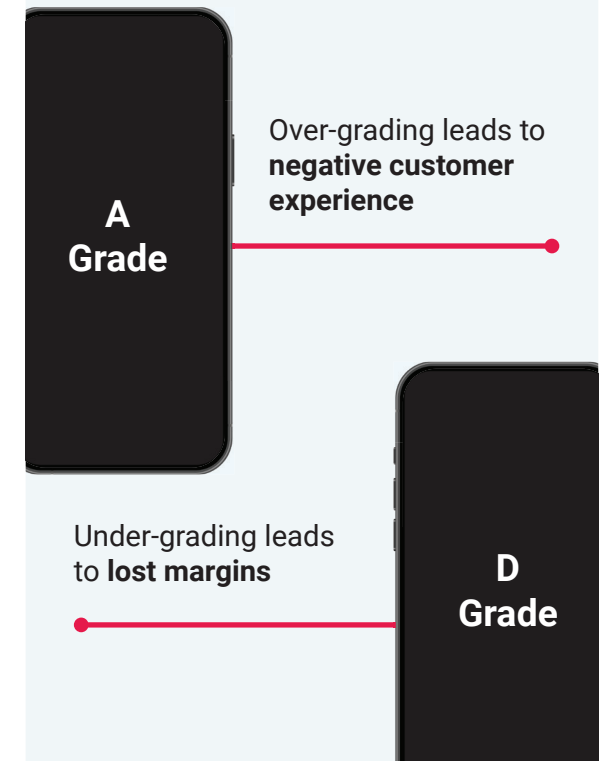
While functionality is a determining factor of a device's resale value, its cosmetic condition—similar to the exterior condition of a used car—is the most significant pricing factor. Scratches, dents, and other cosmetic issues can drastically reduce a device's worth, even if it operates flawlessly.

Significant revenue is at stake. Over-grading a device (the device is not as good as advertised) leads to customer returns and dissatisfaction, causing revenue leakage. Delivering a high-quality consumer experience with a used smartphone is vital. Companies must provide what is expected and minimize returns based on unfulfilled expectations.

Conversely, under-grading (assigning a lower grade than deserved) results in lost profits.

Manual grading is slow and subjective, often reducing the speed and accuracy of the process.

The negative impacts of over-grading and under-grading



The Role of Robotics in the Cosmetic Grading Process

Most companies processing pre-owned smartphones today detect defects and assess the devices' cosmetic condition using a combination of humans and automation. While they value the speed and precision automation can provide, they also understand humans are required to inform the grading algorithm, monitor the robotic processing automation, and focus on device exceptions—those devices that cannot be processed using a fully automated solution.

It's not humans or machines; it's humans *and* machines.

Robotic automation does not need human guidance for defect detection—it will always be exceptional. Robotics delivers precision regarding the number of cosmetic, functional, and structural defects. The robotics findings can be trusted.

Humans help robotics convert the number of defects detected into the easy-to-understand concept of a cosmetic grade.

Apkudo deploys an innovative process to align its customers' device grading criteria with new processing automation by combining human insights with robotic precision. This method overcomes the limitations of subjective defect detection, improves grading efficiency, and ensures accurate device grade assignments.



The Apkudo Cosmetic Grading Alignment Method

Device processing facilities have existing, well-defined cosmetic grading standards used to assess each processed device. Rather than expecting a new customer to adopt Apkudo's automation grading standards, Apkudo works with each customer's team to inform the robotic processing automation using the customer's grading standards and machine learning.

The most efficient way to implement a customer's grading standards is to assess their current defect detection and grading processes, moderate a joint discussion to understand variances, and optimize the automated grading process using larger device volumes.

Assessing a customer's existing defect detection and grading processes

Apkudo evaluates a new customer's current grading standards using existing and trusted employees who have received relevant training and are intimately familiar with detecting

defects to assign the appropriate grade according to the customer's grading standards.

The customer provides a device control group comprising a representative set of devices with varying cosmetic grades.

Utilizing this set of devices, Apkudo leads the customer's graders through:

1. An assessment of a device control group requires graders to use existing grading standards and processes to independently review each control group device. At least three graders assess each device. The duration may vary by customer, ranging from time-limited to unlimited.
2. An Apkudo-moderated and in-depth discussion allows the team to reach a consensus where the same device has received different grades from each grader, identify the correct cosmetic grade according to the number of defects detected and documented grading standards, and understand why human graders have assigned grades outside the documented standards.

The resulting information is used to align Apkudo's grading algorithm and is validated by testing larger volumes of devices.

Key Terms

Accuracy is the the degree to which the result of a measurement, calculation, or specification conforms to the correct value or standard.

Agreement is the harmony or accord in opinion or feeling; a position or result of agreeing.

Controlled expert inspection of a device control group

Apkudo evaluates the results of a controlled expert inspection by assessing the accuracy and agreement of devices' defects detected and corresponding assigned grades.

Apkudo uses the customer device control group, each with a defined cosmetic grade, to document the potential variability and subjectivity associated with the manual evaluation process.

Apkudo wants to document the best humans can do. Even when time is unlimited during the controlled inspection step, the variability between graders persists at the same level. If Apkudo mirrored the production environment—processing hundreds of devices per day with a short amount of time for assessment—human subjectivity would inject bias into the baseline dataset for training the algorithm.

Generally, the process and conditions for manual grading in an actual production environment are strict and aim to achieve consistency. For example, instructions might include:

- Visual inspection is performed at arm's length and under controlled lighting conditions.
- Graders have from four to six seconds to look at each and every device surface.
- Magnification is allowed only for the verification of a defect size.
- The use of tools and gauges is encouraged to aid in acceptance decisions.
- Graders cannot confer with others during the assessment process.

Based on the number of defects observed by the grader, and following the customer's device grading scale, each device is assigned a grade from A to D, with A being the highest/best grade and D being the lowest/worst grade.

The grading scale assumes devices have one distinct grade aligned with defect counts.



However, when working through the controlled expert assessment with customers, Apkudo finds that a device could be assigned multiple grades depending on the grader. (Figure 1)

Apkudo then evaluates the graders' reports using two primary criteria:

- **Accuracy:** Based on the customer's existing grading standards, are graders accurately assessing devices and assigning the correct grades? Not always. Some customers believe their accuracy is 99.95%, but Apkudo has found that the percentage is typically less than 70%.
- **Agreement:** Agreement means different people will evaluate the same device and assign the same grade. Again, based on actual results from testing the grading process with customers, graders only agree approximately 50% of the time.

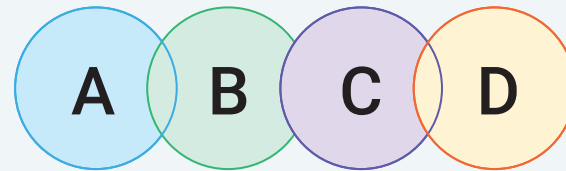
Figure 1

Grading Standards and Outcomes



Customer's Documented Grading Standard

Customer's cosmetic grading standards have clear definitions for each device grade.



Manual Grading by Employees

Human subjectivity often results in the same device receiving two different grades: Human 1 assigns Grade A and Human 2 assigns Grade B.



Automated Grading with Robotics

Robotic processing automation requires clear definitions for each device grade, and each device only receives the correct grade.



Apkudo-moderated joint discussion

Following the evaluation of the device control group, Apkudo moderates a joint discussion with graders to align cosmetic grades and reasoning.

Grade A

The known cosmetic grade for a device from the control group

Accuracy:

60-70% of assessments result in the device receiving Grade A

Agreement:

50% of the grading team agreed the device should be assigned Grade A and not Grade B

The discussion defines and normalizes the terminology used to describe the defects: What delineates a scratch, ding, dent, chip, or crack?

Next, the team determines the number of defects associated with each device grade and how variances (e.g., the length or depth of a scratch) in those defects influence the assigned device grade. This collaborative process helps train robotic systems to understand human intent and establish consistent defect classifications when determining a device's assigned cosmetic grade.

The assessment of the device control group and joint discussion provide the most accurate representation of the manual grading process. This baseline is important because it is an input for the robotic grading algorithms.

Customers investing in automation expect accuracy. Apkudo's extensive work across the industry shows that even the best-in-class graders have less than 70% manual assessment accuracy. Achieving 95% accuracy is impossible when the reference data set from human assessments has an error rate of 30% or higher.

Cosmetic Grading: Scaling Precision

Automation can quickly assess large volumes of devices.

The machine images the device and identifies all the defects and anomalies on all surface sides. It counts and classifies the defects, such as deep scratches, minor dings, and wide cracks. Using an objective robotic assessment, this classification results in the associated grade based on the customer's standards.



Keys to Success

Apkudo's controlled expert inspection and moderated discussion capture the grade intent, but this process demonstrates the high variability of human assessments even in the most controlled conditions. By design, robotics are highly repeatable. Once aligned with the customer's grade intent, the grading algorithm converts machine vision into a device grade determination with higher accuracy and a repeatability rate of 95% or greater, compared to best-in-class human graders of 60%—a 35% increase.

Key Term

The **law of large numbers** is a fundamental principle in probability and statistics that describes how the average of a large number of independent random samples converges to the true value as the sample size increases.

Apkudo leverages the law of large numbers to collect data from a very large set of human-assessed devices.

The law of large numbers averages the variability of human-assessed grades. It provides a population profile that our advanced analytics use to refine the grading algorithm and perform cosmetic tests at scale.

Apkudo works with customers to run large samples of manually graded reference devices across all grades. This step in the process introduces a diversity of surfaces and defects. Groups of variation or device classes are identified—those that require additional measures and confirmation of grading intent. The algorithm adjusts each time it receives new data. Ultimately, the goal is to reduce the variance between people and machines.

If an algorithm is trained too closely to a particular dataset (i.e., overfit), and then a new device is assessed using that algorithm, it is possible the grade will not be accurate. For example, two devices—Device 1 and Device 2—have identical defect profiles, but manual graders have errantly assigned and confirmed different grades for the two devices.



Device 1 has 20 minor defects on the top left quadrant of the rear surface and is assigned a Grade C. Device 2 has 20 minor defects on the lower right quadrant of the rear surface and is assigned a Grade B. The customer grading standard has no grade adjustment for defects by quadrant. The devices must have the same grade according to an objective numerical grading standard guideline. While it is possible to train the robotic system to match the differing manual grades, true accuracy decreases by attempting to achieve identical results based on one correct and one incorrect grade.

In other words, one could perfect the grading for a specific group of devices, but it will no longer represent the true intent of the grading standard in a real production environment. To avoid this, Apkudo utilizes numerical techniques to capture the best available information from the human expert graders and large samples from real-world production environments. See Figure 2 for an example of this linear regression.



The manual grades (blue dots) represent the customer's input for training the robotics system. If the algorithm is fitted to those data points and needs to be 100% accurate with the training data, the training input generates a curve like the red line in the figure.

The green line represents the ground truth—an accurate result. The challenge is that Apkudo does not have the ground truth, but by analyzing many data points, Apkudo can infer the underlying intent of the green line.

The green line represents what Apkudo should be modeling against, and understanding the blue-point offsets represent error and uncertainty in the manual assessment process.

The goal is for robotics to reflect adherence to a standard and make accurate assessments faster. The machine can only be as precise as humans dictate.

The alignment process demonstrates the risk of relying on limited manual inputs as the baseline.

Key Terms

Linear regression is a statistical method utilizing labeled datasets to find an optimized linear function for predicting outcomes on new data. It establishes a relationship between known and unknown variables through a linear equation.

Ground truth is the accurate, real-world data a model aims to predict or replicate. It is the correct or known information used to train and evaluate a model through a labeled dataset.



Conclusion

Cosmetic grading through automation is both an art and a science.

The science lies in robotic defect detection—precise, repeatable, and far more accurate than the human eye. It ensures every imperfection is identified with consistency and reliability. The art comes in translating this raw data into a meaningful cosmetic grade—one that aligns with human expectations and market standards. By blending robotic precision with human insight, automation is redefining the secondary device market. We are committed to partnering with customers to harness this powerful fusion of art and science, driving better business performance and transforming the future of device grading.

Inconsistent human-based cosmetic grading puts customer trust and millions in revenue at risk. Robotic automation eliminates this uncertainty, delivering precise, repeatable assessments that enhance buyer and seller confidence.

As the secondary smartphone market expands, businesses adopting automation will gain a critical competitive edge.

Robotic grading is no longer a future concept—it's a proven, scalable solution that accelerates processing, reduces disputes, and maximizes resale value. Companies integrating robotic precision with human oversight achieve greater accuracy, efficiency, and market credibility. Apkudo's approach ensures grading reflects customer intent while optimizing operational speed and consistency.

The future of secondary device grading is not only about technology—it's about growth. Businesses that embrace automation will unlock new revenue opportunities, strengthen customer trust, and set a new industry standard.

Now is the time to evolve—robotics is improving grading and transforming the secondary device market.



AUTHORED BY:



Gary Landsberg
Chief of Technology Development
Apkudo



Asa Gismervik
Director, Quality Management
Apkudo

About Apkudo

Apkudo helps customers optimize the management of connected devices, ensuring they extract maximum value, reduce costs, and minimize environmental impact. Its Circular Industry Platform offers comprehensive tools for enterprise asset recovery, automated testing and grading, lifecycle management, and seamless resale market integration. By streamlining processes and providing real-time decision support, Apkudo enables businesses to answer the critical question: “What should I do with this device, right now?” Whether assessing value, determining next steps, or facilitating reuse, Apkudo empowers customers to make smarter, more efficient, and sustainable decisions, transforming how they manage device lifecycles.

Copyright © 2025 Apkudo. All rights reserved.
Apkudo and its logo are registered trademarks of Apkudo.



