Amazon Rekognition Face Matching

AWS AI Service Cards

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS AI Service Cards: Amazon Rekognition Face Matching

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Amazon Rekognition Face Matching 1
Overview 1
Intended use cases and limitations 2
Design of Amazon Rekognition Face Matching 3
Deployment and performance optimization best practices 6
Further information 7
Glossary 8

Amazon Rekognition Face Matching

An AWS AI Service Card explains the use cases for which the service is intended, how machine learning (ML) is used by the service, and key considerations in the responsible design and use of the service. A Service Card will evolve as AWS receives customer feedback, and as the service progresses through its lifecycle. AWS recommends that customers assess the performance of any AI service on their own content for each use case they need to solve. For more information, please see AWS Responsible Use of AI Guide and the references at the end. Please also be sure to review the AWS Responsible AI Policy, AWS Acceptable Use Policy, and AWS Service Terms for the services you plan to use.

This Service Card applies to the release of Amazon Rekognition face matching that is current as of October 30, 2024.

Overview

Amazon Rekognition face matching enables application builders to measure the similarity between an image of one face and an image of a second face. This AI Service Card describes considerations for responsibly matching faces in typical identification-style photos and in media (e.g., movies, photo albums and "wild" images captured in uncontrolled or natural environments) using our CompareFaces and SearchFaces APIs. Typically, customers use CompareFaces for comparing a source face with a target face (1:1 matching) and SearchFaces for comparing a source face with a collection of target faces (1:N matching). Amazon Rekognition does not provide customers with pre-built collections of faces; customers must create and populate their own face collections. Throughout this Card, we will use "face matching" to refer to Amazon Rekognition's CompareFaces API and SearchFaces API.

A pair of face images is said to be a "true match" if both images contain the face of the same person, and a "true non-match" otherwise. Given an input pair of "source" and "target" images, Amazon Rekognition returns a score for the similarity of the source face in the source image with the target face in the target image. The minimum similarity score is 0, implying very little similarity, and the maximum is 100, implying very high similarity. Amazon Rekognition itself does not independently decide that two faces from images are a true match or true non-match; the customer's workflow calling CompareFaces and/or SearchFaces decides by using automated logic (by setting a similarity threshold between 0 and 100 and predicting a true match if the similarity score exceeds the threshold), human judgment, or a mix of both.

Human faces differ physically, such as by skin tone and geometry. However, any single individual can be represented by dissimilar images, and, conversely, different individuals may be represented by very similar images. For example, two individuals who only differ in the shape of their eyes might look the same if wearing the same pair of sunglasses. This is because there are many possible factors (called “confounding variation”) that combine to change the location and color of the image pixels that represent a face. These confounding factors include (1) distributions of lighting direction, intensity, and wavelength; (2) head pose; (3) camera focus and imaging defects; (4) pixel resolution; (5) occlusions by hands, facial hair, head hair, cell phones, protruding tongues, scarves, eyeglasses, hats, jewelry, or other objects; (6) facial expression (such as neutral or open-eyed); and (7) alterations to the skin’s tone (for example, by makeup, face paint, sunburn, or acne). Amazon Rekognition's similarity score is designed to be low for images of faces of different individuals and high for images of the same face, ignoring the confounding variations. Amazon Rekognition only uses the information available in the source and target images to assess the similarity of human face images.

Intended use cases and limitations

Amazon Rekognition face matching is only intended to compare faces of humans. It does not support the recognition of faces from cartoons, animated characters, or nonhuman entities. It also does not support the use of face images that are too blurry and grainy for the face to be recognized by a human, or that have large portions of the face occluded by hair, hands, and other objects. Additionally, AWS has implemented a moratorium on police use of the `Rekognition::CompareFaces` and `Rekognition::SearchFaces` API as part of criminal investigations (see section 50.9 of the AWS Service Terms; for more information).

Amazon Rekognition face matching enables many applications, such as identifying missing children, granting access to buildings or conference hospitality suites, verifying identity online, and organizing personal photo libraries. These applications vary by the number of individuals involved, the number of different images available for each individual, the amount of confounding variation expected, the relative costs of false matches and false non-matches, and other factors. We organize these applications into two broad use cases.

- **Identity verification use case:** Identity verification applications use face matching to onboard new users and grant existing users access to resources. In this use case, confounding variation is usually minimized by using photos from government issued IDs (such as passports and drivers licenses) and real-time selfies that encourage front-facing poses of well-lit, unobscured faces. This allows each individual in the target collection to be represented by a small number of face

images and allows the number of different individuals in the collection to be large (such as in the millions). In this use case, certain end users might try to fool the system to gain access, so customers may mitigate this risk by, for example, manually checking that the source and target images submitted to Amazon Rekognition meet the customer's expectations, and/or requiring matches to have high similarity scores (for example, 95).

- **Media use case:** Media applications use face matching to identify individuals in photos and videos from a set of known individuals (for example, finding family members in vacation videos). In this use case, there is high confounding variation between source and target images of the same individual, so target collections might contain fewer individuals with more images per user (perhaps spanning multiple years of the person's life). There is less incentive for end users to try to fool the system in this use case, so customers may elect to have highly automated workflows, and, given the high confounding variation, may allow matches to have lower similarity scores (for example, 80).

Design of Amazon Rekognition Face Matching

Machine learning

Amazon Rekognition face matching is built using ML and computer vision technologies. It works as follows: (1) Locate the portion of an input image that contains the face. (2) Extract the image region containing the head, and align the region so the face is in a "normal" vertical position, outputting cropped face images. (3) Convert each cropped face image to a "face vector" (technically, a mathematical representation of the image of a face). Note that the collections searched by SearchFaces are sets of face vectors, not sets of face images. (4) Compare the source and target face vectors and return the system's similarity score for the face vectors. For more information, see Overview of face detection and face comparison in the *Amazon Rekognition Developer Guide*.

Performance expectations

Individual and confounding variation will differ between customer applications. This means that performance will also differ between applications, even if they support the same use case. Consider two Identity Verification applications A and B. With each, a user first enrolls their identity with a passport-style image, and later verifies their identity using real-time selfies. Application A enables smartphone access by using the smartphone camera to capture selfies that are well-lit, well-focused, frontally-posed, high resolution, and unoccluded. Application

B enables building access by using a doorway camera to capture selfies that are less well lit, blurrier, and lower resolution. Because A and B have differing kinds of inputs, they will likely have differing face matching error rates, even assuming that each application is deployed perfectly using Amazon Rekognition.

Test-driven methodology

We use multiple datasets to evaluate performance. No single evaluation dataset provides an absolute picture of performance. That's because evaluation datasets vary based on their demographic makeup (the number and type of defined groups), the amount of confounding variation (quality of content, fit for purpose), the types and quality of labels available, and other factors. We measure Amazon Rekognition performance by testing it on evaluation datasets containing pairs of images of the same individual (matching pairs), and pairs of images of different individuals (non-matching pairs). We choose a similarity threshold, use Amazon Rekognition to compute the similarity score of each pair, and based on the threshold, determine if the pair is a match or a non-match. Overall performance on a dataset is represented by two numbers: the true match rate (the percentage of matching pairs with a similarity above the threshold) and the true non-match rate (the percentage of non-matching pairs with a similarity score below the threshold). Changing the similarity threshold changes the true match and true non-match rates. Groups in a dataset can be defined by demographic attributes (e.g., gender), confounding variables (e.g., the presence or absence of facial hair), or a mix of the two. Different evaluation datasets vary across these and other factors. Because of this, the true match and non-match rates – both overall and for groups – vary from dataset to dataset. Taking this variation into account, our development process examines Amazon Rekognition's performance using multiple evaluation datasets, takes steps to increase true match and/or true non-match rates for groups on which Amazon Rekognition performed least well, works to improve the suite of evaluation datasets, and then iterates.

Fairness and bias

Our goal is that Amazon Rekognition face matching work well for all human faces. To achieve this, we use the iterative development process described above. As part of the process, we build datasets that capture a diverse range of human facial features and skin tones under a wide range of confounding variation. We routinely test across use cases on datasets of face images for which we have reliable demographic labels such as gender, age, and skin tone. We find that Amazon Rekognition performs well across demographic attributes. For example, iBeta, a NIST-accredited lab, performed a third-party evaluation of Amazon Rekognition. The evaluation used an Identity Verification dataset containing high-quality images with good lighting conditions, no blur, and no occlusion. The dataset was split across six demographic groups based on skin

tone and gender. iBeta observed at a similarity threshold set to 95, the lowest true match rate seen across all six demographic groups was 99.97185% and the lowest true non-match rate seen across all six demographic groups was 99.99988%. Because performance results depend on a variety of factors including Amazon Rekognition, the customer workflow, and the evaluation dataset, we recommend that customers do additional testing of Amazon Rekognition using their own content.

Explainability

If customers have questions about the similarity score returned by Amazon Rekognition for a given pair of source and target images, we recommend that customers use the bounding box and face landmark information returned by Amazon Rekognition to inspect the face images directly.

Robustness

We maximize robustness with a number of techniques, including using large training datasets that capture many kinds of variation across many individuals. Because Amazon Rekognition cannot simultaneously have very high sensitivity to small differences between different individuals (such as identical twins) and have very low sensitivity to confounding changes (such as makeup applied to enhance cheekbones), customers must establish expectations for true match and true non-match rates that are appropriate to their use case, and test workflow performance, including their choice of similarity threshold, on their content.

Privacy and security

Amazon Rekognition face matching processes three kinds of data: customer input images, face vectors of input images, and output similarity scores and output metadata. Face vectors are never included in the output returned by the service. Inputs and outputs are never shared between customers. Customers can opt out of training on customer content via AWS Organizations or other opt out mechanisms we may provide. For more information, see Section 50.3 of the AWS Service Terms and the AWS Data Privacy FAQs. For service-specific privacy and security information, see Amazon Rekognition FAQs and Amazon Rekognition Security.

Transparency

Where appropriate for their use case, customers who incorporate Amazon Rekognition face matching APIs in their workflows should consider disclosing their use of ML and face recognition technology to end users and other individuals impacted by the application, and give their end users the ability to provide feedback to improve workflows. In their documentation, customers can also reference this AI Service Card.

Governance

We have rigorous methodologies to build our AWS AI services responsibly, including a working backwards product development process that incorporates Responsible AI at the design phase, design consultations and implementation assessments by dedicated Responsible AI science and data experts, routine testing, reviews with customers, and best practice development, dissemination, and training.

Deployment and performance optimization best practices

We encourage customers to build and operate their applications responsibly, as described in the AWS Responsible Use of AI Guide. This includes implementing Responsible AI practices to address key dimensions including fairness and bias, robustness, explainability, privacy and security, transparency, and governance.

Workflow Design: The accuracy of any application using Amazon Rekognition face matching depends on the design of the customer workflow, including: (1) the number of unique individuals being matched, (2) the amount of confounding variation allowed, (3) selection of similarity thresholds, (4) how matches are decided, (5) how consistently the workflow is applied across demographic groups, and (6) periodic retesting for drift.

1. **Individual variation:** When searching for a source face among a collection of target faces, success increases with the degree of physical dissimilarity between the different individuals in the target set. For example, matching between identical twins is substantially harder than matching between fraternal twins or unrelated individuals. In general, target collections with larger numbers of unique individuals pose a higher risk of having two unique individuals who appear similar, and require more care when making a final decision about a match. Workflows should consider the possible similarity of individuals in the target collection when interpreting the similarity scores returned for source images.
2. **Confounding variation:** When selecting pairs of source and target images, workflows should include steps to minimize variations between the source and target images (such as differences in lighting conditions). If variation is high, consider adding multiple face images (“options”) for each target individual that cover the expected variations (such as poses, lighting, and ages), and comparing the source face image with each target option. If it is only practical to have a single option, consider using a passport-style, front-facing, unoccluded headshot. Workflows should establish policies for permissible input images, and monitor compliance by periodically and randomly sampling inputs.

3. **Similarity thresholding:** It is important to set an appropriate similarity threshold for the application. Otherwise, the workflow could conclude that there is a match where there is not (a false match) or vice versa (a false non-match). The cost of a false match may not be the same as the cost of a false non-match. For example, an appropriate similarity threshold for authentication might be much higher than that for media. To set an appropriate similarity threshold, a customer should collect a representative set of input pairs, label each as a match or non-match, and try higher or lower similarity thresholds until satisfied.
4. **Human oversight:** If a customer's application workflow involves a high risk or sensitive use case, such as a decision that impacts an individual's rights or access to essential services, human review should be incorporated into the application workflow where appropriate. Face matching systems can serve as tools to reduce the effort incurred by fully manual solutions, and to allow humans to expeditiously review and assess possible matches and non-matches.
5. **Consistency:** Customers should set and enforce policies for the kinds of source and target images permitted, and for how humans combine the use of similarity thresholding and their own judgment to determine matches. These policies should be consistent across all demographic groups. Inconsistently modifying source and target images or similarity thresholds could result in unfair outcomes for different demographic groups.
6. **Performance drift:** A change in the kinds of images that a customer submits to Amazon Rekognition, or a change to the service, may lead to different outputs. To address these changes, customers should consider periodically retesting the performance of Amazon Rekognition, and adjusting their workflow if necessary.

Further information

- For service documentation, see CompareFaces, SearchFaces.
- For an example of an authentication workflow design, see Identity Verification Using Amazon Rekognition.
- For details on privacy and other legal considerations, see the following AWS policies: Acceptable Use, Responsible AI, Legal, Compliance, and Privacy.
- For help optimizing workflows, see Generative AI Innovation Center, AWS Customer Support, AWS Professional Services, Ground Truth Plus, and Amazon Augmented AI.
- If you have any questions or feedback about AWS AI service cards, please complete this form.

Glossary

Controllability: Steering and monitoring AI system behavior.

Privacy & Security: Appropriately obtaining, using and protecting data and models.

Safety: Preventing harmful system output and misuse.

Fairness: Considering impacts on different groups of stakeholders.

Explainability: Understanding and evaluating system outputs.

Veracity & Robustness: Achieving correct system outputs, even with unexpected or adversarial inputs.

Transparency: Enabling stakeholders to make informed choices about their engagement with an AI system.

Governance: Incorporating best practices into the AI supply chain, including providers and deployers.