

Detekcja ataków kaszlu i zestresowanego głosu

Funkcja analiz audio dostępna w wybranych kamerach
z wbudowanym mikrofonem

Maj 2024

Streszczenie

Detekcja ataków kaszlu i zestresowanego głosu to aplikacja do analiz audio, która wykrywa incydenty dzięki ciągłemu nasłuchowi otoczenia.

Aplikacja ta jest fabrycznie instalowana w wybranych kamerach Axis z wbudowanym mikrofonem. Składa się z dwóch odrębnych algorytmów detekcji, które mogą być używane osobno lub łącznie.

Detektor kaszlu wykrywa pojedyncze kasznięcia lub ataki kaszlu, umożliwiając personelowi szybkie reagowanie na potrzeby osób wymagających pomocy. Detektor zestresowanego głosu identyfikuje wzorce dźwiękowe związane z przymusem, złością lub strachem, dzięki czemu jest skutecznym narzędziem zapobiegania przestępstwom, ograniczania napaści lub wskazywania osób potrzebujących pomocy.

Jeśli chodzi o prywatność, jest ona chroniona, ponieważ prawidłowe działanie *detekcji ataków kaszlu i zestresowanego głosu* nie wymaga zapisywania danych audio. Jeśli rejestrowanie materiału audio nie zostało świadomie włączone, żadne dźwięki nie są rejestrowane.

Konfigurując kilka ustawień, można zapewnić optymalne działanie analiz w konkretnym zastosowaniu. Ponadto funkcja *detekcji ataków kaszlu i zestresowanego głosu* wykonuje stałe kontrole stanu, aby zapewnić prawidłowe działanie.

Spis treści

1	Wprowadzenie	4
2	Detektor kaszlu i detektor zestresowanego głosu	4
3	Umożliwienie optymalnej audio detekcji	5
4	Świadomość wieloczuJNIkowa	5
5	Nakładki	6
6	Typy zdarzeń i kondycja	6
7	Prywatność	7

1 Wprowadzenie

Dostępne w kamerze funkcje analityczne z audio detekcją są zaawansowanym uzupełnieniem dozoru wizyjnego. Umożliwiają wczesną detekcję i ostrzeżenie o potencjalnych incydentach, często przed ich wykryciem w materiale wizyjnym.

Detekcja ataków kaszlu i zestresowanego głosu to aplikacja do analiz audio, która non stop nasłuchuje otoczenie oraz klasyfikuje i filtruje dźwięki. W przypadku wykrycia ataku kaszlu lub zestresowanego głosu aplikacja generuje alert.

W tym dokumencie omówiono funkcję *detekcji ataków kaszlu i zestresowanego głosu* oraz sposób jej skonfigurowania pod kątem optymalnej detekcji.

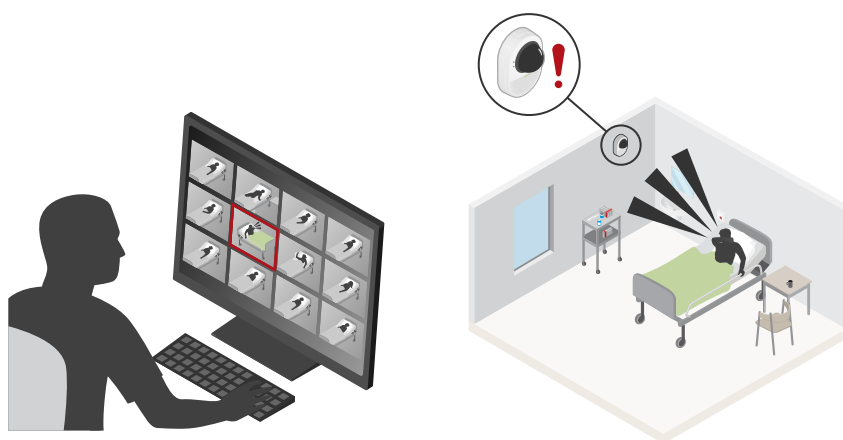
2 Detektor kaszlu i detektor zestresowanego głosu

Detekcja ataków kaszlu i zestresowanego głosu to funkcja fabrycznie instalowana w wybranych kamerach Axis z wbudowanym mikrofonem. Detektory na bieżąco, bezpośrednio w kamerze wychwytyją dźwięki sygnalizujące incydenty. Można się zdecydować na używanie jednego z detektorów lub obu jednocześnie.

Działanie detektora kaszlu polega na wykrywaniu kaszlnięć i ich liczeniu w określonym przedziale czasu. Umożliwia to personelowi szybką reakcję w sytuacji, gdy ktoś zacznie kasłać lub dostanie ataku kaszlu. W zależności od konfiguracji detektor może wykrywać ataki kaszlu lub pojedyncze kaszlnięcia.

Aby ograniczyć występowanie wielu powiadomień o zdarzeniach w bardzo krótkim czasie, po policzeniu pierwszego kaszlnięcia rozpoczyna się 5-sekundowy czas blokady. Jeśli funkcja analityczna jest skonfigurowana na 3 kaszlnięcia w ciągu 30 sekund, policzy następne kaszlnięcie tylko wtedy, gdy od poprzedniego upłynie co najmniej 5 sekund. Ewentualne kaszlnięcia podczas 5-sekundowego czasu blokady nie zostaną policzone. Oznacza to, że przy opisanych ustawieniach alert zostanie wysłany dopiero po zliczeniu 3 kaszlnięć, między którymi upłynęło co najmniej 5 sekund.

Detektor zestresowanego głosu identyfikuje wzorce dźwiękowe związane z przymusem, złością lub strachem w głosie monitorowanej osoby. Po rozpoznaniu takiego wzorca system wysyła automatyczne powiadomienie do personelu w postaci alertu wizualnego lub przez wyzwolenie alarmu. Wczesne ostrzeżenie umożliwia personelowi szybką reakcję. Personel może pomóc osobie znajdującej się w potrzebie lub zapobiec eskalacji, która w przeciwnym razie mogłaby doprowadzić do agresji fizycznej.



Zastosowanie detekcji ataków kaszlu i zestresowanego głosu w ochronie zdrowia.

3 Umożliwienie optymalnej audio detekcji

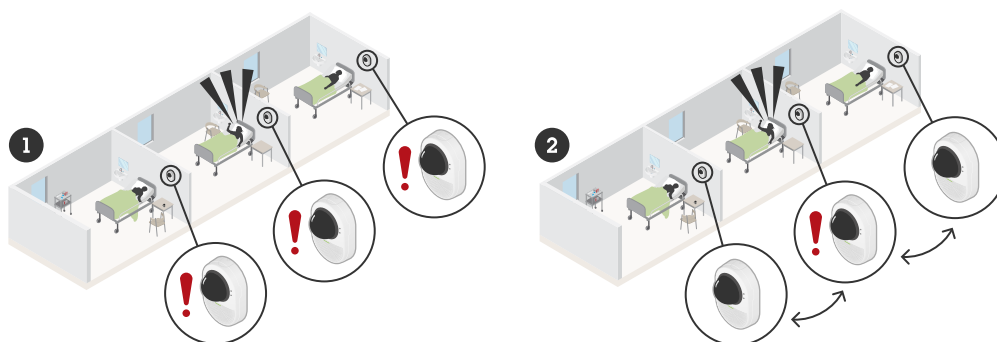
- **Lokalizacja kamer.** Kamerę z funkcją analityczną należy umieścić minimum 1,5 metra od źródeł zakłócającego hałasu, takich jak systemy ogrzewania, wentylacji i klimatyzacji, systemy nagłośnieniowe lub głośniki oraz trzaskające drzwi. Ponadto kamerę najlepiej jest zlokalizować w linii widzenia obszaru, który ma być objęty audio detekcją. Chociaż istnienie linii widzenia nie jest ścisłym wymogiem, może zwiększyć dokładność detekcji. Wynika to z faktu, że ewentualne narożniki lub przeszkody znajdujące się na drodze dźwięku mogą go osłabić. Przykładowo fale dźwiękowe o różnych częstotliwościach uginają się pod różnymi kątami.
- **Czułość.** System detekcji można precyzyjnie dostroić przy użyciu ustawień czułości. Wyższa czułość przekłada się na większą liczbę detekcji. Zwiększa to ryzyko zbędnych detekcji (fałszywych alarmów), ale może być wymagane w przypadkach, gdy każde zdarzenie musi zostać wykryte. Przy niższej czułości detekcja będzie zgłaszana tylko w przypadku, gdy istnieje pewność co do prawidłowego sklasyfikowania dźwięku. Zwiększa to ryzyko przegapienia potencjalnych incydentów, ale niska czułość może być wymagana, jeśli w przeciwnym razie byłoby zbyt wiele fałszywych alarmów.
- **Tryb zbierania danych.** Na pewien czas po instalacji można włączyć tryb zbierania danych, aby bliżej poznać rodzaje wykrywanych dźwięków. Przeprowadzona analiza i jej wyniki mogą pomóc w określeniu optymalnego poziomu czułości w danej instalacji.
- **Próg detekcji kaszlu.** Można ustawić wartość progową określającą wymaganą liczbę kaszlnięć. Jeśli liczba kaszlnięć osiągnie wartość progową we wskazanym okresie, spowoduje to wyzwolenie alarmu.
- **Ustawienia zaawansowane.** Ustawienia zaawansowane są przeznaczone tylko dla doświadczonych użytkowników. Zmiana ustawień może prowadzić do nieprawidłowych detekcji lub ich zupełnego braku. Jednak w pewnych scenariuszach może wystąpić potrzeba zmodyfikowania tych ustawień. Modyfikacji należy dokonywać tylko na polecenie eksperta systemowego lub po zasięgnięciu jego opinii.

4 Świadomość wieloczujnikowa

Kiedy kamery są umieszczone blisko siebie, na przykład w sąsiadujących pomieszczeniach, ten sam incydent audio może zostać wykryty przez większą liczbę kamer. To może utrudnić dokładne ustalenie jego lokalizacji.

Jeśli chodzi o detekcję zestresowanego głosu, w takich przypadkach może pomóc funkcja świadomości wieloczujnikowej. Gdy jest ona aktywna i wiele kamer wykryje ten sam głos zestresowanej osoby, powiadomienie wzbudzi wyłącznie kamera, która wykryła go jako pierwsza. W ten sposób sąsiadujące

kamery współpracują, aby ograniczyć liczbę fałszywych zdarzeń i zduplikowanych powiadomień o tym samym zdarzeniu.



- 1 *Brak świadomości wieloczujnikowej: kamery umieszczone w sąsiadujących pomieszczeniach wykrywają ten sam zestresowany głos i wyzwalają wiele alarmów.*
- 2 *Dzięki świadomości wieloczujnikowej detekcję zgłasza tylko najbliższa położona kamera.*

Funkcja świadomości wieloczujnikowej pozwala tworzyć grupy kamer równorzędnych, które obejmują kamery z pokrywającymi się zasięgami detekcji dźwięku. Istnieją pewne ograniczenia:

- We wszystkich kamerach równorzędnych należy skonfigurować korzystanie z synchronizacji czasu NTP.
- Wszystkie kamery równorzędne powinny korzystać z tej samej wersji funkcji *detekcji ataków kaszlu i zestresowanego głosu*.
- Wszystkie kamery równorzędne powinny mieć możliwość wzajemnego kontaktowania się za pośrednictwem sieci.

Jeśli któryś z powyższych warunków nie zostanie spełniony, kamera równorzędna przełączy się w tryb autonomiczny i zgłosi stan obniżonej wydajności.

5 Nakładki

Na przekaz wizyjny można nałożyć aktualizowany na żywo spektrogram i powiadomienia aplikacji. Nakładki umożliwiają zmianę rozmiaru, a także zmianę położenia przez przeciągnięcie w żądane miejsce. Odrębny suwak pozwala dostosować przezroczystość nakładki.

Powiadomienia aplikacji pokazują zdarzenia wykryte przez kamerę i aktualny stan aplikacji.

Spektrogram przedstawia wizualną reprezentację dźwięku. Słyszac dźwięk i jednocześnie widząc jego reprezentację graficzną, w wielu przypadkach można szybciej ustalić wagę incydentu.

6 Typy zdarzeń i kondycja

Zdarzenia generowane przez detektor kaszlu i detektor zestresowanego głosu mają charakter *bezstanowy*. Są one chwilowymi wystąpieniami wyzwalanymi przez detekcję. Gdy upłynie czas bloku zdarzenia (pięciosekundowy, z możliwością skonfigurowania), ewentualna detekcja generuje nowe zdarzenie.

Kondycję funkcji *detekcji ataków kaszlu i zestresowanego głosu* odzwierciedlają *zdarzenia o charakterze stanowym*. Oznacza to, że określony stan zdarzenia pozostaje aktywny tak długo, jak długo występuje dany stan faktyczny, i zmienia się dopiero po jego ustąpieniu.

Wbudowane kontrole kondycji umożliwiają weryfikację prawidłowego działania i sygnalizowanie ewentualnych nieprawidłowości. W przypadku funkcji *detekcji ataków kaszlu i zestresowanego głosu* można wyróżnić trzy stany kondycji:

- Stan prawidłowy: normalne działanie. Jest możliwa detekcja.
- Stan obniżonej wydajności: funkcja działa w trybie obniżonej wydajności. Zazwyczaj jego przyczyną jest czynnik tymczasowy, taki jak utrata powiązanej kamery, obcinanie dźwięku ze względu na jego bardzo dużą głośność lub przepełnienie bufora audio. W stanie obniżonej wydajności detekcja jest możliwa, ale liczba fałszywych lub nieudanych detekcji może być większa. Stan obniżonej wydajności zazwyczaj ustępuje samoistnie.
- Stan awarii: brak działania. Detekcja jest niemożliwa. Zazwyczaj przyczyną tego stanu jest czynnik, który nie ustępuje samoistnie, taki jak wyłączenie obsługi audio w ustawieniach urządzenia lub wyciszenie wejścia audio.

Stany obniżonej wydajności i awarii są sygnalizowane w panelu informacyjnym i w nakładce tekstowej (jeśli jest włączona), dzięki czemu operator wie, że aplikacja działa z obniżoną wydajnością lub wykryła awarię.

Co 60 sekund jest wyzwalane zdarzenie pulsu (jeśli odpowiednia funkcja jest włączona; częstotliwość pulsu można zmodyfikować). Dzięki niemu po stronie odbiorczej można sprawdzić, czy funkcja analityczna działa, a w razie nieodebrania pulsu wyzwoić alarm. W stanie awarii zdarzenia pulsu nie są wysyłane.

7 Prywatność

Dane audio są przetwarzane i analizowane w kamerze, a prawidłowe działanie *detekcji ataków kaszlu i zestresowanego głosu* nie wymaga zapisywania żadnych danych dźwiękowych. Podczas zdarzeń dźwięk jest rejestrowany tylko w przypadku świadomego włączenia funkcji rejestrowania. Ta opcja może się przydać podczas dochodzeń wyjaśniających przyczyny incydentów, rozwiązywania problemów w przypadku zgłoszenia fałszywych trafień lub odsłuchiwanie incydentów, jeśli nie umożliwia tego system zarządzania materiałem wizyjnym.

O firmie Axis Communications

Axis umożliwia tworzenie mądrzejszego i bezpieczniejszego świata, tworząc rozwiązania zwiększające bezpieczeństwo i wydajność biznesową. Jako firma z branży technologicznej będąca liderem na rynku, Axis oferuje systemy dozoru wizyjnego, kontroli dostępu, domofonowe i rozwiązania audio. Rozwiązania te są wzbogacone o inteligentne aplikacje analityczne i wysokiej jakości szkolenia

Firma Axis zatrudnia około 4000 zaangażowanych pracowników w ponad 50 krajach i współpracuje z partnerami z sektora technologii oraz integracji systemów na całym świecie, aby dostarczać rozwiązania dla klientów. Firma Axis powstała w 1984 roku, a jej siedziba znajduje się w Lund w Szwecji