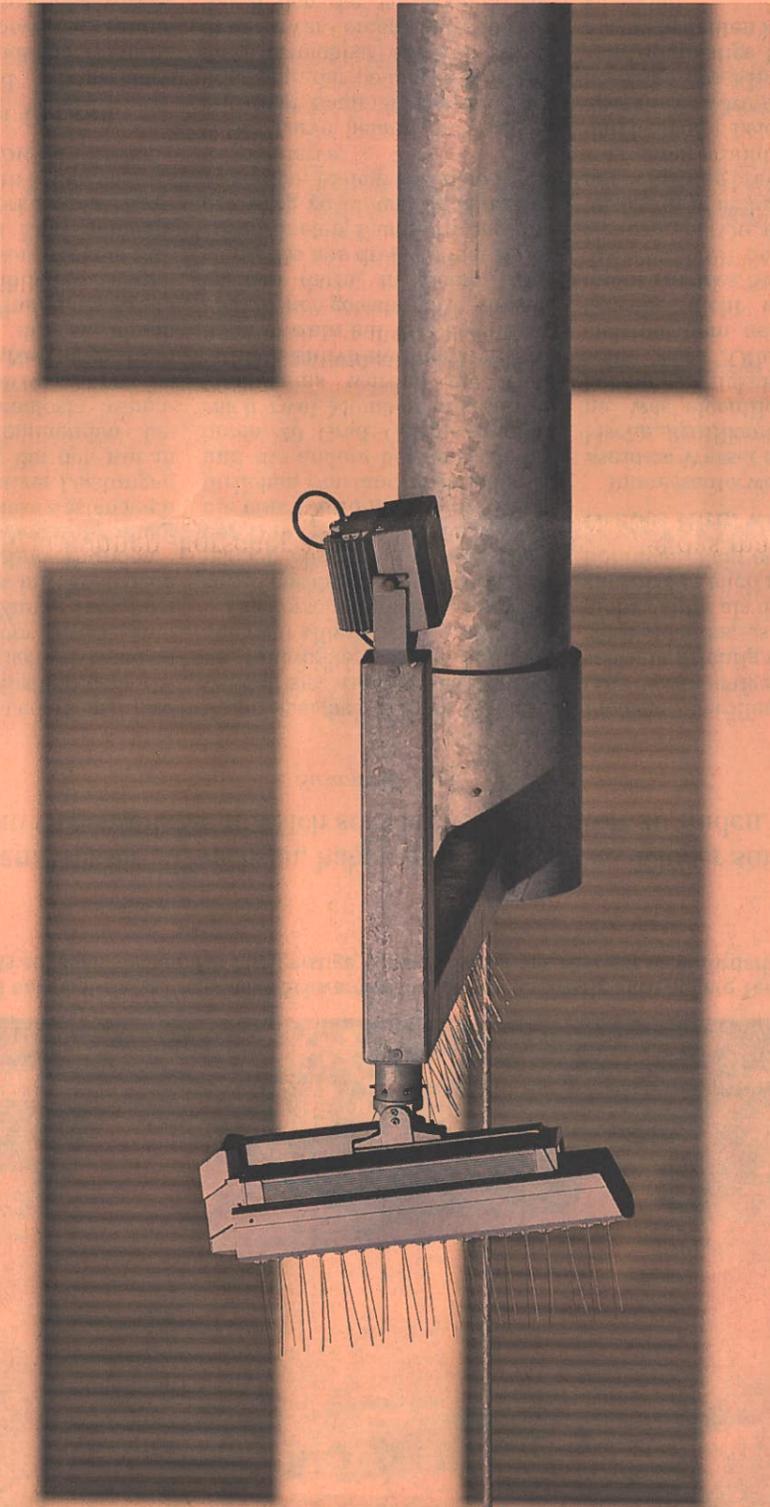


Wie künstliche Intelligenz lernt, Bewegungsmuster zu verstehen

Computersysteme tun sich noch schwer, relevante Informationen aus Videobildern zu extrahieren. Ein Forschungsprojekt zur Videoüberwachung brachte Verbesserungen.

Alois Punhösel



Für den Objektschutz sind Kameras heute unerlässlich. Die Analyse der Bilder erfolgt zunehmend automatisch. Foto: Imago/Ipom

Zahlreiche Monitore in einem Kammerchen, die Live-Bilder von Kameras vor Hausen- fahrten, auf Gängen und Vorpätzen zeigen; ein Security-Mitarbeiter, der davor sitzt und die Bildschirme mehr oder weniger aufmerksam beobachtet – das ist das klassische Szenario einer Videoüberwachung, wie sie lange Zeit im Ob-

jektschutz von Ministerien, Museen oder Bibliotheken verbreitet war. Dank des Booms in der automatischen Bilderkennung durch künstliche Intelligenz (KI) und Machine-Learning-Algorithmen kommen aber auch hier zunehmend intelligentere Systeme zum Einsatz. Sie schlagen Alarm, wenn die Kamerabilder ungewöhnliche Situationen zeigen.

Der Haken dabei: Der Anteil an Fehlalarmen ist meist noch viel zu hoch. Das liegt daran, dass vielerlei Phänomene auftreten können, die zwar ungewöhnlich sind, aber keineswegs auf verdächtige Vorgänge, unbedingtes Betreten oder Gefahrsituationen hindeuten. Eine Spiegelung in einer Glasfassade, der Stumm, der Blätter durchs Bild weht, eine Spinne, die sich vor der Kamerallinse abseilt – ungeklärte Bewegungen dieser Art sind für die Systeme kaum einzuordnen. Das führt dazu, dass das Sicherheitspersonal alle diese Kamerabilder nun erst wieder von Hand kontrollieren muss.

„Schutz der Außenhaut“

Im Projekt „Skin“ haben sich Experten in den letzten Jahren einem verbesserten „Schutz der Außenhaut kritischer Infrastrukturen“ gewidmet. Gemeinsam mit dem Sicherheitstechnikdienstleister PKI, bei dem die Projektleitung lag, und dem Innenministerium, dem Verteidigungsministerium und dem Vienna Centre for Social Security (VICSS) als Partner arbeiteten Forscher des Software Competence Center Hagenberg (SCCH) daran, die automatisierten Analyse-systeme verlässlicher zu machen und Fehlalarme zu reduzieren. „Skin“ wurde im Rahmen des Sicherheitsforschungsprogramms Kiras unterstützt, das im Verantwortungsbereich des Landeswirtschafsministeriums angesiedelt ist und von der Förderagentur FFG gemanagt wird.

Der Schlüssel zur Verbesserung der Überwachungssysteme ist für Bernhard Moser, den Forschungsdirektor des SCCH, die Information, die in den Bewegungsabläufen steckt – im Gegensatz zur Analyse jeweils einzelner Bilder. „Der Mensch hat genug Fantasie, um sich ausgehend von einem statischen Bild eine Szene vorzustellen und die Bewegung, den Kontext mitzudenken“, sagt Moser. „Der künstlichen Intelligenz fehlt diese Fähigkeit.“

Im Rahmen von Behavioral Analytics, die auf zeitliche Abfolgen in Kamerabildern fokussiert, wird das „Verständnis“ des Computers für eine Situation erhöht. Moser: „Wenn jemand Kaffee holt, hat das einen bestimmten Ablauf. Die Bewegungen haben eine innere Logik, eine Grammatik – ähnlich der eines Satzes.“ Kein Wunder also, dass die KI-Systeme, die die Bewegungsmuster untersuchen, jenen der Sprachanalyse ähneln.

Die Systeme sollen etwa aufgrund einer An-fangsbeobachtung Vorhersagen treffen. Wenige Momente später wird verglichen, ob die Realität der nachkommenden Bilder mit diesen Prognosen übereinstimmt. Der Autschlüsselung der Kamerabilder liegt dabei eine Unzahl von statistischen Berechnungen zugrunde. Objekte werden über mehrere Bilder hinweg getrackt. „Für den Menschen ist klar, dass eine Person, die hinter einer Säule verschwindet, wahrscheinlich auf der anderen Seite gleich wieder hervor kommt – für den Computer aber keineswegs“, sagt Moser. Das System erkennt eine Vielzahl sogenannter Tracklets, Bruchstücke von Bewegungsabläufen, um aufgrund dieser auf Trajektorien, also längerfristige Bewegungsabläufe, zu schließen. Letztendlich soll dank statistischer Analysen eine Einschätzung entstehen, welche Aktionen häufig vorkommen und was eine Anomalie sein könnte.

Anomalien-Cluster

Die Forscher haben das System mit Videoaufzeichnungen aus einem ganzen Jahr hinweg trainiert, um der KI beizubringen, welche Situationen üblich und „normal“ sind. Anomalien werden nicht einfach nur gemeldet, sondern in Clustern mit gleichartigen Vorgängen zusammenfasst. Diese werden durch Menschen kontrolliert – wobei der Computer die Information erhält, ob sie aus Sicherheitsperspektive relevant sind oder nicht. Die Einschätzungen werden so mit der Zeit genauer. Auf diese Art lernt die KI, dass der Cluster mit den sich vor der Kamera abspielenden Spinnen vielleicht nicht relevant ist, veranschaulicht Moser. Andere Ereignisse – etwa wenn eine Person über eine Mauer steigt – sollen als berechnete Alarmansätze verbleiben.

Insgesamt konnte laut den Ergebnissen von „Skin“ die Fehleranfälligkeit der Videoüberwachung um etwa 65 Prozent pro Kamera und Tag reduziert werden. Die Ergebnisse sind Grundlage einer Reihe weiterer Projekte: etwa zur Verhaltensvorhersage im Sicherheitsbereich – zum Beispiel an Flughäfen –, aber auch zur Optimierung des Stadtgüterverkehrs oder zur Überwachung des toten Winkels bei Schienenfahrzeugen.

