

Capítulo 4

Pruebas y resultados.

Los análisis experimentales se llevaron a cabo para explorar el poder discriminativo de los diccionarios en la segmentación de humo de las secuencias de video al aire libre.

Se utilizó un conjunto de datos de 6 secuencias de video tomadas por las cámaras de vigilancia con diferentes longitudes y condiciones en términos de luminosidad con diferentes escenarios o incidentes de fuego.

Los conjuntos de datos de entrenamiento fueron extraídos de los primeros fotogramas de las secuencias de fragmento. En la mayoría de los casos cámaras vigilan vastas áreas y están situadas lejos del incendio forestal, por lo tanto las secuencias de video probadas se caracterizan por el hecho de que el humo aparece en el paisaje.

Cada diccionario es entrenado para cada región (humo, cielo y tierra), sin embargo hay una gran diversidad de objetos en las regiones de tierra (como lo es la vegetación, colinas, llanuras, estructuras artificiales y montañas) que dependen de la ubicación de la cámara.

En lugar de entrenar un diccionario grande con representación esparsiva de las regiones de terreno, el diccionario aprende de forma independiente para cada cámara evitando la necesidad de un gran número de átomos logrando reducir el tiempo de segmentación.

Antes del aprendizaje los diccionarios son inicializados por una extracción automática y aleatoria de átomos de cuadros delimitadores trazados manualmente que contienen la clase de interés. La información de los átomos se generaran a través de parches de imagen de tamaño 3x3.

El número total de entradas por átomo es $N=99$, que incluye 27 valores de intensidad de los tres planos RGB contenidas en el parche, 54 valores de color de intensidad adicionales de parches en el anterior y fotograma siguiente que supone la información temporal (cuboide espacio-temporal) y la información estadística (media y varianza) para cada plano de color y cada ubicación temporal que asciende a 18 entradas adicionales.

Después de la inicialización los diccionarios se entrenan de una manera supervisada mediante la construcción de conjuntos de vectores característica de entrenamiento con un proceso de extracción guiada similarmente a la de un diccionario de inicialización.

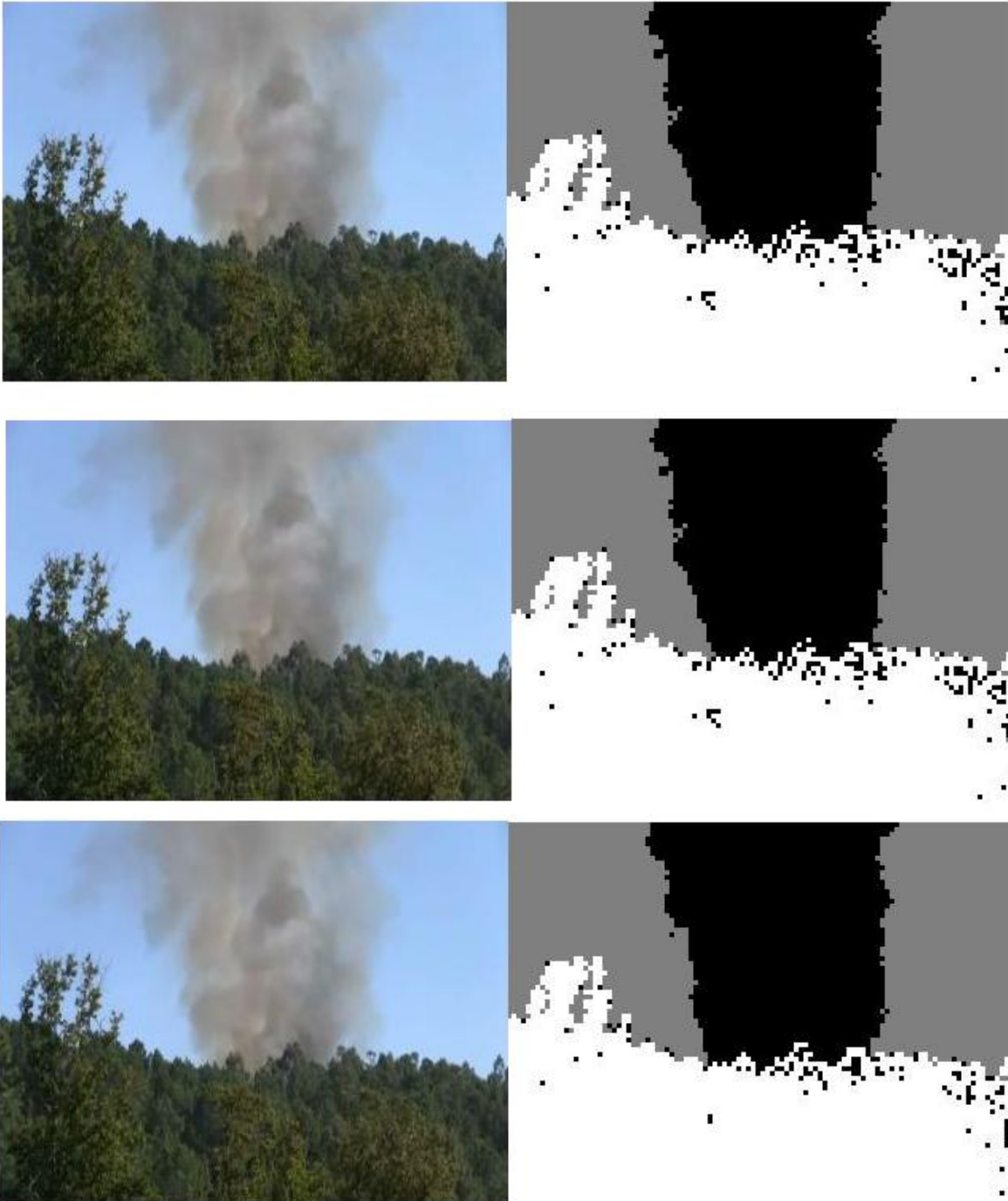


Figura 4 Segmentación de incendio forestal.

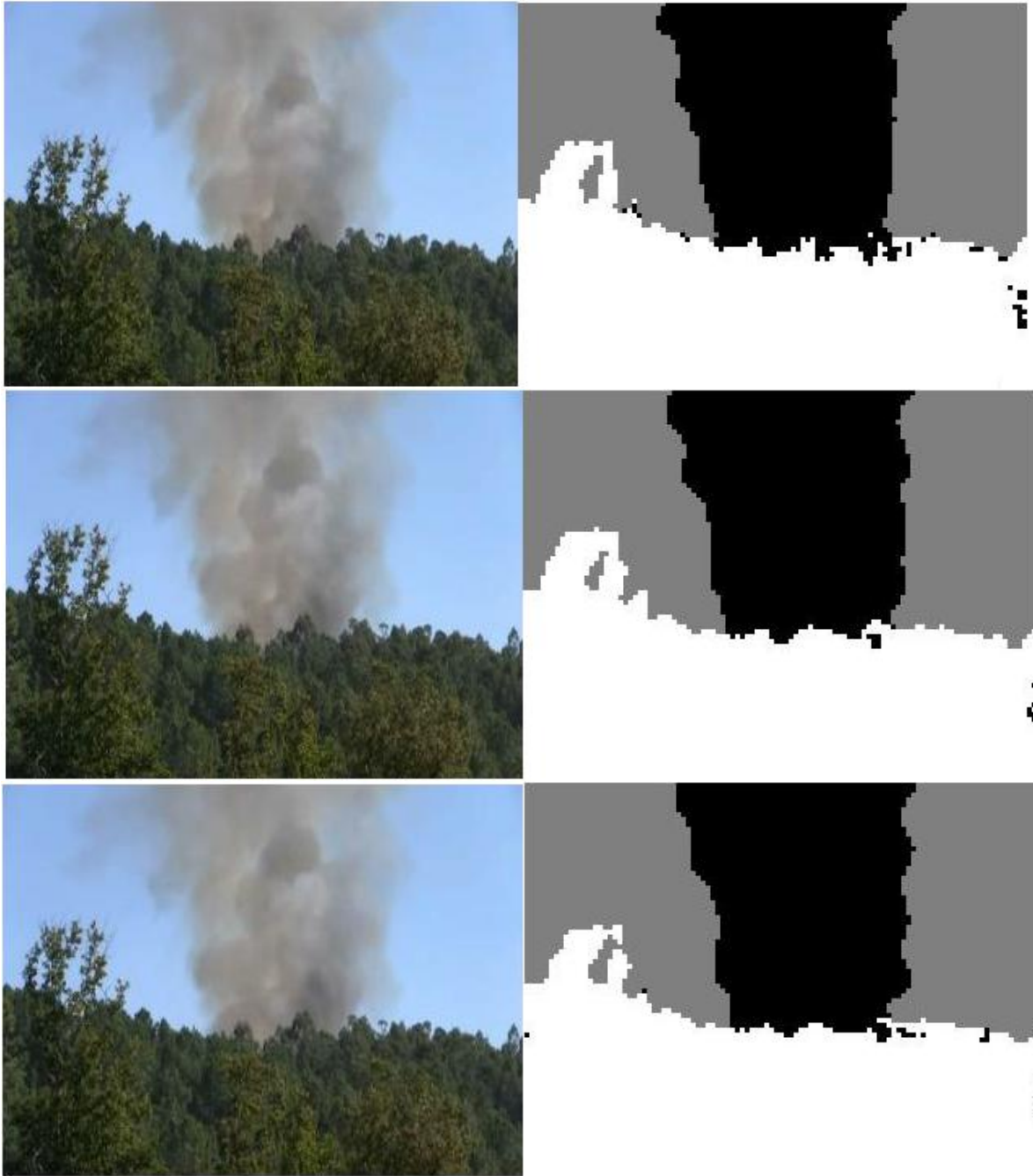


Figura 5 Reclasificación con cuboide espacio-temporal.

La figura 4 y figura 5 muestra un fragmento corto de la segmentación de una secuencia de código correspondiente a un evento de fuego. El tiempo de separación entre frames es de 1seg y el tamaño de cada imagen es de 326x 484 pixeles.

Esta prueba se realizó con 200 átomos por diccionario y 120 vectores de característica de entrenamiento para los diccionarios de aprendizaje supervisado con 90 iteraciones del método K-SVD y un factor de esparcion $L=3$.

El tiempo computacional medio en el aprendizaje para un diccionario supervisado es de 34,5 seg en una aplicación de MATLAB 2010a con un procesador Intel Core 2 Duo a 2,8GHz. El tiempo computacional promedio durante la segmentación de una trama es de 107 seg.

Para reducir el tiempo de segmentación se eliminaron aquellos átomos que no fueron ajustados durante el aprendizaje. El tiempo computacional promedio durante la segmentación de un cuadro con los diccionarios reducidos es de 81,3 seg.

Los parches clasificados como parte de humo, cielo o regiones de tierra se muestran en negro, gris y blanco respectivamente. Las manchas de humo mal clasificadas (falsos positivos) son aquellas manchas en el interior de las regiones blancas y grises. Para reducir la presencia de parches mal clasificados en un segmento se utilizó un cuboide espacio-temporal, donde el parche central del cuboide se reclasifico con una votación mayoritaria utilizando los parches vecinos en el interior del cuboide.

Después de la segmentación y la reclasificación se utilizó un algoritmo adicional de verificación para realizar un seguimiento de los cambios de áreas en regiones de interés (ROI) [12].

Si la ROI es cada vez mayor sobre los frames siguientes entonces se le confirma como una región de humo ya que se considera que las regiones de humo crecen.

La búsqueda para la ROI en el primer fragmento de frame se lleva a cabo en toda la trama a través de una exploración desde los 4 bordes hacia el centro hasta que se encuentren los bordes del cuadro delimitador más pequeño que encierra nuestra ROI.

Para el caso de los frames siguientes, el cuadro limite está determinado por una aproximación basada en el cuadro limite de los frames anteriores. El seguimiento de los cambios de área en regiones de interés ayuda a reducir las falsas alarmas.

La figura 6 muestra los resultados de detección de humo en dos fotogramas de una secuencia de frames del fuego controlado antes y después del inicio del fuego con un tiempo transcurrido de 4 minutos entre ambos marcos.

Esta prueba se realizó con 480 átomos por diccionario y 96 vectores característica de entrenamiento. El tiempo medio calculado es de 33 segundos para el aprendizaje de diccionario supervisado. Se utiliza un factor esparsivo de $L=3$, y un tiempo medio de de 93 segundos para la segmentación de cada fragmento de frame.

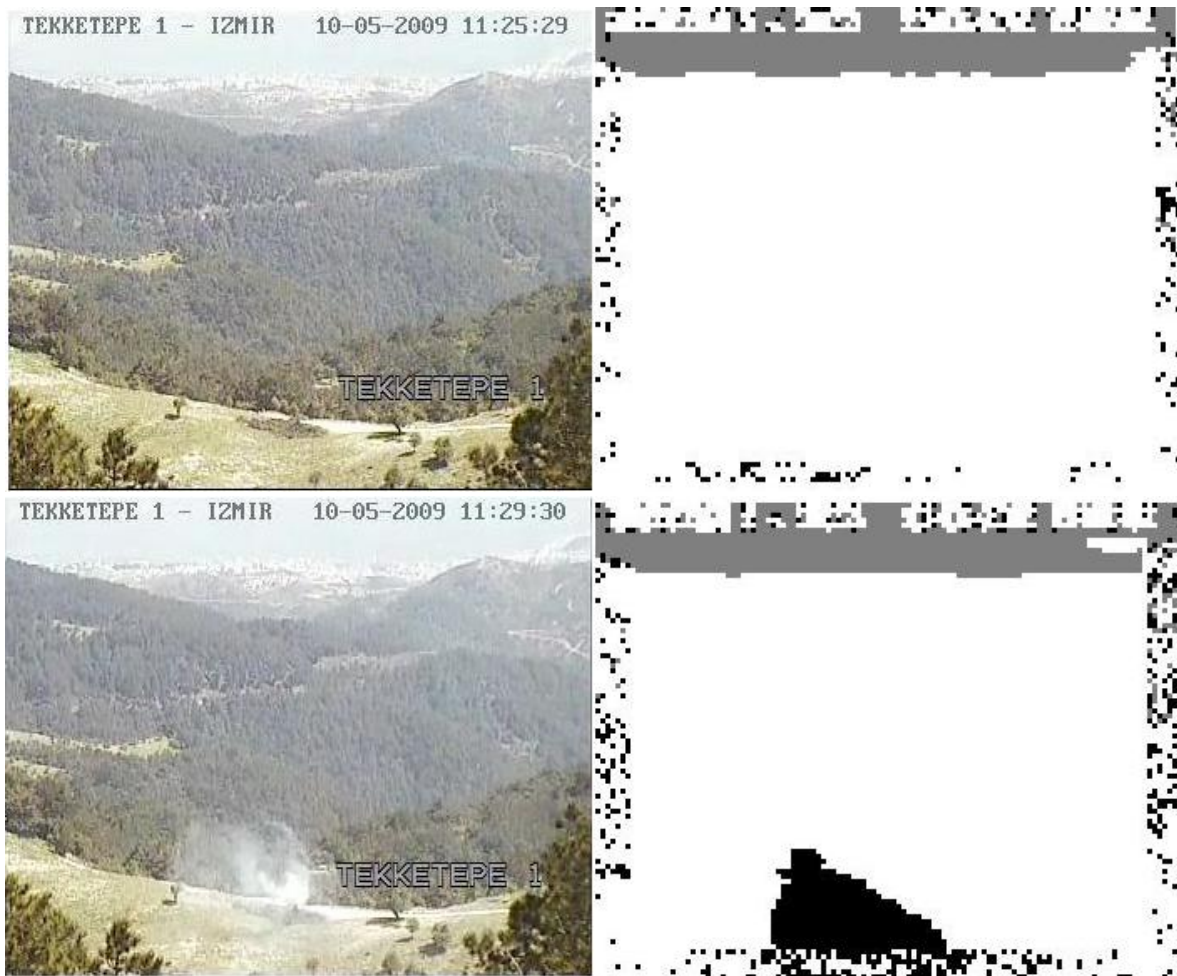


Figura 6 Prueba de detección de dos frames en un fragmento de secuencia antes y después del fuego.

El algoritmo de detección fue probado siendo útil en secuencias de fragmento de código con frames bajo condiciones que hacen que su segmentación sea una tarea fácil, tales como alta iluminación en el medio ambiente como se muestra en la figura 7; o escenarios con estructuras hechas por el hombre tal como se muestra en la figura 8.



Figura 7 Detección de humo a alta iluminación

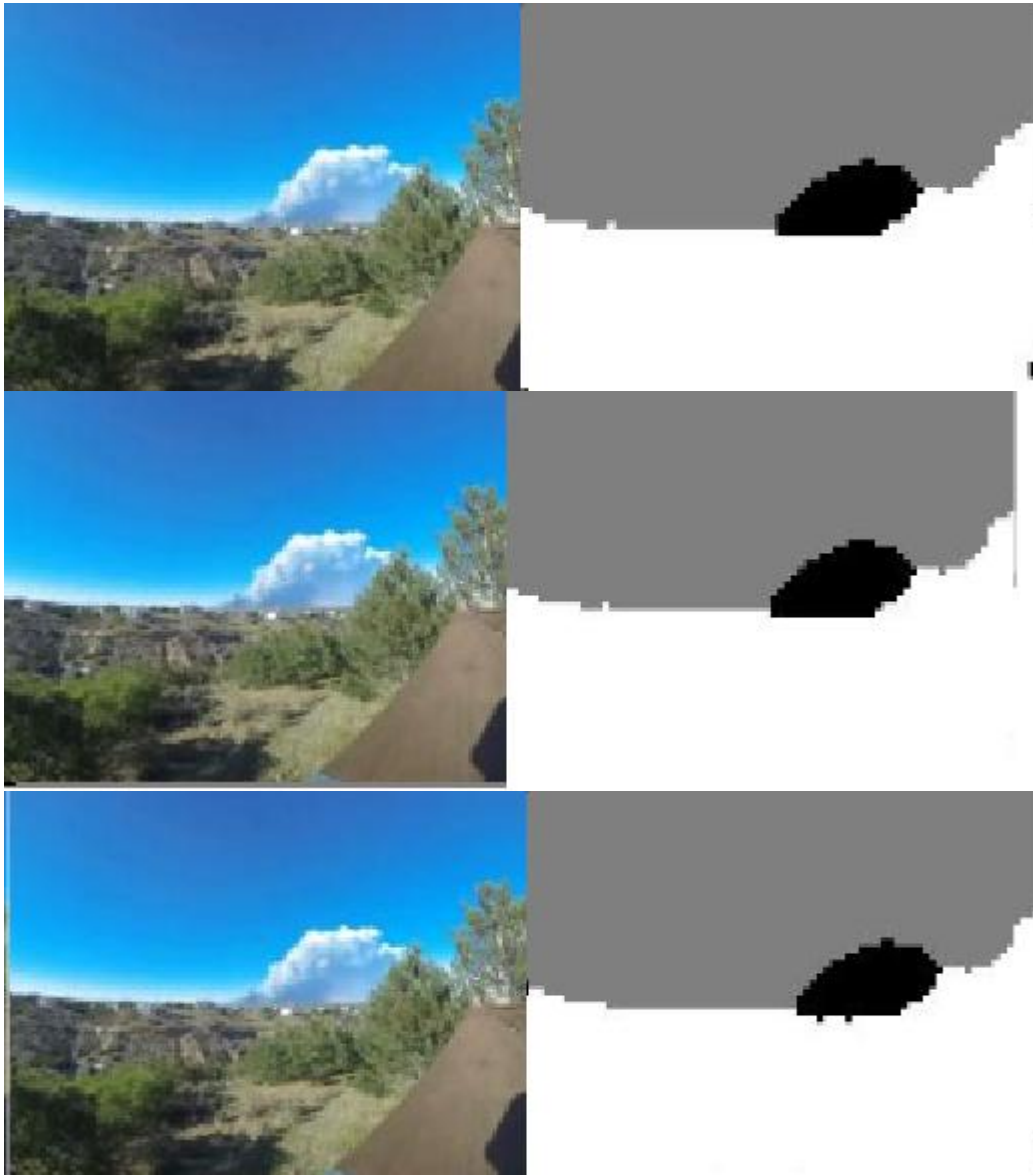


Figura 8 Detección del humo en un escenario con una estructura hecha por el hombre.



Figura 9 Detección de humo tomando en cuenta la presencia de nubes

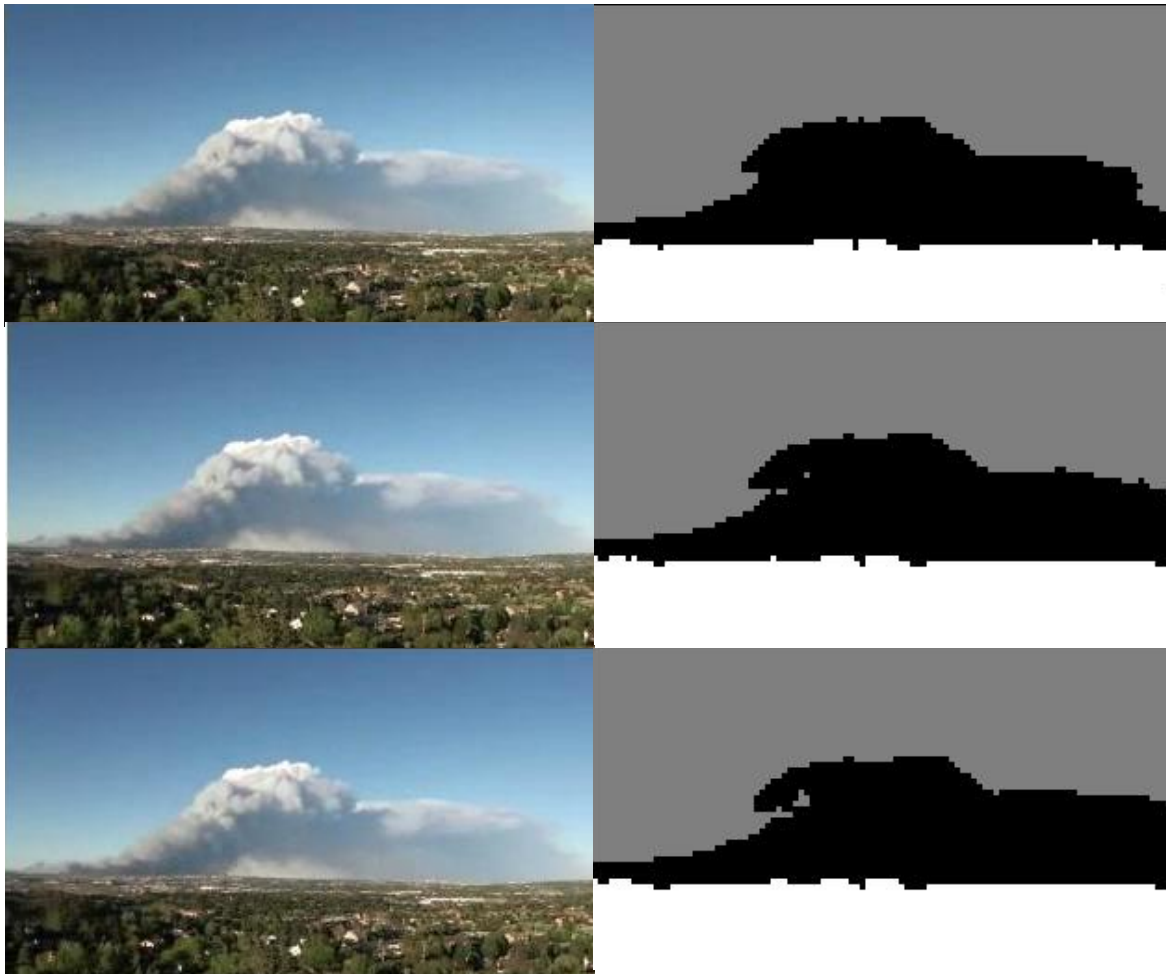


Figura 10 Cámara con movimiento.

El rendimiento de detección se evaluó mediante el cálculo de la tasa de detección (DR) y la tasa de las falsas alarmas (FAR). Esos frames utilizados para el entrenamiento, extraídos de las secuencias no se utilizan para medir el desempeño.

La tasa de detección se obtiene considerando todos los marcos de diferentes secuencias de código que muestrean un incidente de fuego, y esta tasa se define como:

$$DR = \frac{\text{Number of Frames with True Positives}}{\text{Number of Frames with True Positives} + \text{Number of Frames with False Negatives}}, \quad (27)$$

Donde el número de frames con falsos negativos es cero, lo que da una tasa de reconocimiento del 100 %. El verdadero desafío que enfrenta el algoritmo de detección de humo, emerge con la tasa de alarmas definido como:

$$FAR = \frac{\text{Number of Frames with False Positives}}{\text{Number of Frames with False Positives} + \text{Number of Frames with True Negatives}}, \quad (28)$$

Donde el número de frames con falsos negativos se reduce considerablemente después de que el uso de la interacción espacio-temporal entre los parches vecinos da una tasa de alarma de 8 %. Dando seguimiento de los cambios de áreas en regiones de interés reduciendo la tasa de falsas alarmas hasta el 5.7 % a expensas de introducir una latencia crucial.

Método	DR	FAR
B. Chul Ko et.al [13]	86.5 %	No reportado
L. Wang et.el [14]	No reportado	No reportado
A.Ravichandran et.el [15]	No reportado	No reportado
V.Vipin [16]	99 %	14 %
R. González et. el [2]	100 %	0 %
Método propuesto	100 %	5.7 %

Tabla 1 Porcentaje de Tasa de detección y Tasa de falsas alarmas.







Snippets	Frames	TP	FN	TN	FP
Snippet 1 28 frames		28	0	0	0
Snippet 2 66 frames	<small>LENGUETZE 3 - 12M13 30-05-2009 11:29:30</small> 	33	0	33	2
Snippet 3 30 frames		30	0	0	0
Snippet 4 30 frames		30	0	0	0
Snippet 5 10 frames		10	0	0	0
Snippet 6 30 frames		30	0	0	0

Tabla 2 Tasa de detección y tasa de falsas alarmas en el que se aplico el algoritmo propuesto.