

SINTONIZACION DE HÍPER PARÁMETROS DE SVM PARA DETECCION DE PEATONES

Ruiz Varela Oscar Ramsés, González Rojo Sergio A., Navarro Gómez Carmen Julia, Gómez Parra Ana Luisa

Instituto Tecnológico de Chihuahua

Instituto Tecnológico de Chihuahua

Ave. Tecnológico #2909, Chihuahua, Chih. México. C.P. 31310

+51 (614) 2-01-2000

Universidad Autónoma de Chihuahua

Servicios Educativos del Estado de Chihuahua

orruiz@itch.edu.mx, grojo@itch.edu.mx, cjnavarro@uach.mx, ana.gomez@seech.edu.mx

RESUMEN.

La detección de peatones es una actividad con renovado interés por la comunidad que participa en el desarrollo de sistemas ADAS (Advanced Driver Assistance Systems) y PPS (Pedestrian Protection Systems) como componentes de vehículos autónomos. La Aplicación de Redes Neuronales, Máquinas de vector soporte y otras técnicas de inteligencia artificial proporcionan resultados prometedores, pero tienen posibilidades de caer en mínimos locales en lugar de mínimos globales, cuya consecuencia es la obtención de resultados inferiores a los mejores posibles. En el presente trabajo se realizó una implementación para detectar peatones utilizando una combinación de HOG (Histograma de Gradiente Orientado), SVM (Máquina Vectorial de Apoyo), Y se definieron los parámetros (híper) de la SVM que determinan el rendimiento del sistema para el objetivo particular. Se aplicaron algunas técnicas como validación cruzada para encontrar el mejor modelo obtenido por la selección calculada de híper parámetros. Se sugieren algunas estrategias para optimizar la sintonización de sistemas de este tipo.

Palabras Clave: SVM, sintonización de parámetros, detección de peatones.

ABSTRACT.

Pedestrian detection has acquired major interest for community trying to build ADAS and PPS systems as part of passengers autonomous vehicles. Artificial neural networks, support vector machines and related artificial intelligence techniques yields breakthroughs, but could deliver minimum local values instead of minimum global values, resulting in lower outputs than best. Presented work it's a system made of components like HOG (Histogram Oriented Gradient), SVM (Support Vector Machine), etc. System's hyper parameters that defines performance were obtained for the desired goal. Several techniques like cross validation were applied to find best model through calculated hyper parameters selection. Several strategies to optimize parameters tuning of systems like this one are suggested.

Keywords: SVM, parameters tuning, pedestrian detection.

1. INTRODUCCION

Por la intención de desarrollar prototipos en plataformas de bajo desempeño (Raspberry PI) en tiempo real se selecciona utilizar una SVM en lugar de una red neuronal convolucional o alguna otra técnica de mayor potencial, dado que se cumple la tarea en menor tiempo. Misma razón para seleccionar C++ en lugar de Python.

Debido al uso incrementado de aprendizaje de máquina en empresas. La optimización de híper parámetros es de substancial interés comercial, y cada vez representa una función más importante, ya sea como herramientas internas de la compañía, como parte de servicios de nube inteligentes o como un servicio por sí mismo [6].

La detección de peatones en sistemas de transporte autónomo es de vital importancia y un reto por la enorme diversidad de escenarios y condiciones posibles (lluvia, nieve, objetos o material en caminos, postura del peatón, atuendo, etc.) [1]

El Histograma de Gradiente Orientado (HOG) convierte una imagen en una secuencia numérica que permite visualizar el patrón que identifica objetos y formas específicas en dichas imágenes. Dichas secuencias numéricas se definen como rasgos. Se utilizaron imágenes del dataset de INRIA para el entrenamiento de la SVM y para las pruebas sobre los resultados del entrenamiento de la misma. El reto para las SVM y para las técnicas de redes neuronales continúa siendo la selección de parámetros para problemas específicos, es decir, una regla de selección de parámetros óptimos de SVM para detección de peatones no proporciona el mejor resultado si el objetivo es utilizar SVM para predecir brotes de dengue. Otro reto de ambas técnicas es la cantidad de fuentes de datos de las que se dispone y su calidad [3].

2. MAQUINAS SVM.

El mejor desempeño de generalización para una determinada tarea, con una cantidad finita de datos de entrenamiento, se alcanzará si se logra el correcto balance entre la exactitud alcanzada en ese conjunto de entrenamiento en particular y la habilidad de la máquina para aprender de cualquier conjunto de entrenamiento sin error [2].

Alta habilidad de aprendizaje, buena generalización en clasificación y regresión, bajo costo de entrenamiento, coloca a las SVM como buenas opciones de solución a problemas de pronóstico de carga eléctrica, reconocimiento de patrones, construcción de modelos de comportamiento crediticio, etc. [4]. SVM es un algoritmo de aprendizaje supervisado basado en kernels para separar dos o más clases inducidas del conjunto de datos de entrenamiento. La meta es producir un clasificador que predecirá correctamente ejemplos no utilizados por esta durante el entrenamiento, o sea, proporcionar una buena generalización.

Sean m ejemplos de entrenamiento (x_i, y_i) , $y_i = \pm 1$, $i=1,2,3,\dots,m$, existe un hiper plano $w \cdot x + b = 0$ que separa los casos positivos y negativos usando una función de decisión

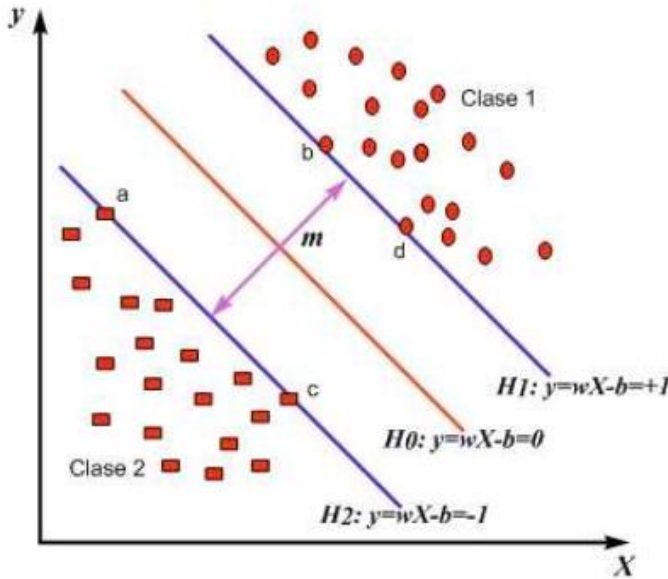
$$f(x) = \text{sign}(w \cdot x + b) \quad (1)$$

Donde

$$\text{sign}(x) = \begin{cases} -1 & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (2)$$

para clasificaciones binarias, donde w es el vector normal al hiper plano, conocido como vector de pesos y b es la desviación o bias [4].

Figura 1. Separación de hiper plano



En la figura 1 existen dos hiper planos y la región entre ellos es la banda de margen, dada por:

$$\text{Max } \frac{2}{\|w\|^2} \quad \text{ó} \quad \text{Min } \frac{\|w\|^2}{2} \quad (3)$$

Sujeta a las restricción $y_i (w \cdot x_i + b) \geq 1$, $\forall i = 1,2,3,\dots,m$

La mayoría de los dataset reales contienen ruido. El uso de un margen suave reduce los efectos del ruido y los elementos aislados. Por lo que se introduce el margen suave con el error marginal ξ_i , la función objetivo es entonces:

$$\text{Min } \frac{\|w\|^2}{2} + \sum_{i=1}^m \xi_i \quad (4)$$

Con

$$y_i (w \cdot x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad \forall i = 1,2,3 \dots m \quad (5)$$

El parámetro C controla el compromiso entre el error marginal y el error de las pruebas ⁴.

Las condiciones de Karush-Kuhn-Tucker (KKT) son necesarias para resolver problemas no lineales óptimamente. El principal problema se convierte en un problema dual y aplicando las condiciones KKT obtenemos:

$$w = \sum_{i=1}^m \alpha_i y_i x_i \quad (6)$$

Usando funciones kernel, la función de decisión se torna en

$$f(x) = \text{sign} \left(\sum_{i \in S^v} \alpha_i y_i K(x, x_i) + b \right) \quad (7)$$

Con la función objetivo:

$$\text{Max } w(\alpha) = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j \alpha_i \alpha_j K(x_j, x_j),$$

$$\forall i = 1 \dots m \quad (8)$$

con

$$\sum \alpha_i y_i = 0 \quad (9)$$

Donde la función $K(x,z)$ definida en $R^d \times R^d$ es llamada un kernel si existe un mapeo ϕ del espacio R^d al espacio Hilbert, $\phi: R^d \rightarrow H$ tal que $K(x,z) = \langle \phi(x), \phi(z) \rangle$.

Las funciones kernel que han sido ampliamente utilizadas desde el inicio de la teoría de SVM son: el kernel de función de base radial (RBF o Gaussiana),

$$K(x, z) = e^{-\gamma \|x-z\|^2} \quad (10)$$

Y el kernel lineal

$$K(x, z) = x^T z \quad (11)$$

En la ecuaciones (8) y (10) puede observarse el papel que los hiper parámetros C y γ juegan en el proceso de optimización cuando la función $K(x,z)$ es el RBF [5].

3. DETECCIÓN DE PEATONES.

La detección de peatones con imágenes puede utilizarse en sistemas de vigilancia de video y sistemas de asistencia para el conductor. Representa un reto mayor que detectar otros objetos rostros y automóviles, porque la apariencia de la gente presenta muchas fluctuaciones tales como vestuario, postura, iluminación [7].

Se extrajeron rasgos HOG de regiones locales de 16×16 . Se calculan gradientes de bordes y su orientación por cada pixel en estas regiones locales. Se utiliza la aproximación de los filtros Sobel para los cálculos de la derivada. La magnitud del gradiente $m(x,y)$ y la orientación $\phi(x,y)$ se calculan de la siguiente manera:

$$m(x, y) = \sqrt{dx(x, y)^2 + dy(x, y)^2} \quad (12)$$

$$\varphi(x, y) = \begin{cases} \tan^{-1}\left(\frac{dx(x, y)}{dy(x, y)}\right) - \pi & \text{si } dx(x, y) < 0 \text{ y } fy(x, y) < 0 \\ \tan^{-1}\left(\frac{dx(x, y)}{dy(x, y)}\right) + \pi & \text{si } dx(x, y) < 0 \text{ y } fy(x, y) > 0 \\ \tan^{-1}\left(\frac{dx(x, y)}{dy(x, y)}\right) & \text{en otro caso} \end{cases} \quad (13)$$

Las regiones locales se dividen en pequeñas áreas espaciales llamadas celdas. El tamaño de las celdas es de 4x4 píxeles. Se calculan histogramas de gradiente de bordes, con 8 orientaciones para cada celda. Cada posible orientación en las celdas de 4x4 se considera un rasgo HOG. Por lo que cada vector HOG contiene 128 rasgos.

4. IMPLEMENTACIÓN.

El desarrollo del software tiene como objetivo realizar la detección de peatones y evaluar el desempeño de la detección variando los hiper de la SVM, buscando una correlación entre el valor de los hiper seleccionados y la tasa de detección de peatones.

La tabla 1 muestra parte de los resultados obtenidos. Cada corrida realizó un nuevo entrenamiento a partir de los parámetros seleccionados. Los resultados se almacenan como renglones en un archivo csv.

Tabla 1. Comparativo de resultados de detección de acuerdo a los parámetros C y gamma seleccionados.

total	positivos	kernel	gamma	C	entrena
100	90	rbf	1	100	300
100	66	rbf	1	10	300
100	93	rbf	1	1	300
100	72	rbf	1	10	300
100	96	rbf	10	1	300
100	71	rbf	100	1	300

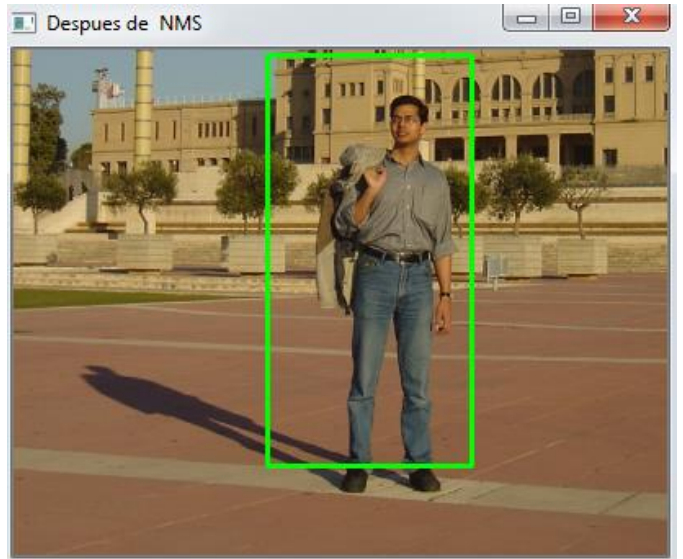
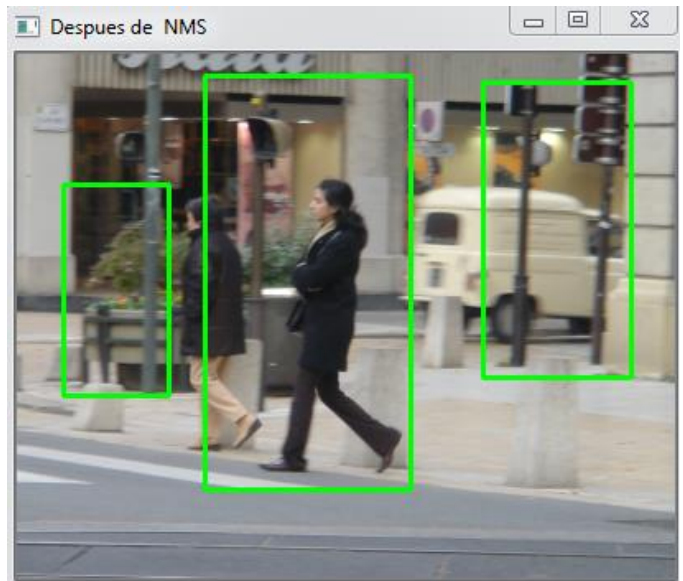
La Tabla 1 permite ver que la mejor selección de hiper se da con los valores C=1 y gamma=10, entrenando con 300 imágenes.

Los resultados más bajos con una detección del 66% ocurre con los hiper C=10 y gamma=1.

El comparativo del mejor y el peor resultado muestra una relación no lineal entre los hiper seleccionados y el resultado.

La Figura 2 nos muestra dos ejemplos de la detección con error (falso positivo) y correcta

Figura 2. Detección de peatones con imágenes del dataset INRIA



5. CONCLUSIONES.

La implementación realizada permite la detección de peatones entrenando y probando con un dataset, aunque una mejor aproximación hubiese sido la validación con al menos otro dataset.

Se observa necesario realizar un análisis del dataset para medir la cantidad de sesgo que se tiene del conjunto de imágenes como un factor a considerar para compensarlo.

Otra actividad que se debe realizar es aplicar un análisis estadístico (Pearson R u otro) para seleccionar los valores de los parámetros de una manera sistemática y cuantificar estadísticamente el impacto de las variaciones en los parámetros respecto al desempeño obtenido.

Referencias.

- [1] D. Geronimo, A.M. Lopez, A.D.Sappa. Survey of pedestrian detection for advanced driver assistance system. Vol 32, No 7, Julio 2010, pag 1239
- [2] C. J. C. Burguess . A tutorial on support vector machines for pattern recognition, data mining and knowledge discovery. (1998) pages 121-167.
- [3] Z Mustaffa, Y. Yusof.. A comparison of normalization techniques in predicting dengue outbreak. 2010 International Conference on Business and economic research. 2011.
- [4] K. S. Parik, T.P. Shah. Support vector machine- a large margin classifier to diagnose skin illnesses. 3rd International Conference on innovations on automation and Mechatronic Engineering, ICIAME 2016.
- [5] L.C. Padierna, A.R. Domínguez, H. J. Puga Soberanes, R. Baltazar. Hyper Parameter tuning for support vector machines by estimation of distribution algorithms. www.researchgate.net. Publication 311555174. Diciembre, 2017.
- [6] M. Feurer, F. Hutter. AutoML. Methods, Systems, Challenges. <https://www.automl.org/book/> .
- [7] T. Kobayashi, A. Hidaka, T. Kurita. Selection of histograms of oriented gradients features for pedestrian detection. International conference of oriented gradient features for pedestrian detection. 2007. Pag 598-607