

Calculo automático de índice de masa corporal usando Visión Artificial

Juan de Jesús Amador Reyes¹, Josué Espejel Cabrera², Farid García Lamont²,
Laura D. Jalili² y José Sergio Ruiz Castilla²

juandjador@gmail.com, jeco309@hotmail.com, fglamont@yahoo.com.mx, lauradojaliz@gmail.com y jsergioruizc@gmail.com

¹ Universidad Autónoma del Estado de México, Campus Ecatepec, José Revueltas No. 17 Col Tierra Blanca, C.P. 55020, Ecatepec, México.

² Universidad Autónoma del Estado de México, Campus Texcoco, Prolongación. Av. Zumpango s/n Fracc. El Tejocote, C.P. 56259, Texcoco, México.

Pages: 454–471

Resumen: Los avances tecnológicos se han implementado en diferentes campos de la vida diaria, un campo de gran relevancia es la medicina. La investigación de metodologías que faciliten la determinación temprana de los síntomas y su tratamiento posterior es de vital importancia. Este artículo propone un método basado en la visión artificial para el cálculo del índice de masa corporal (IMC), para este propósito se reunieron personas de diferentes edades y complejiones, a partir de las imágenes obtenidas de todas las personas se obtienen las características del área de interés para obtener un buen desempeño al clasificar el IMC, obteniendo una precisión de clasificación del 73%.

Palabras-clave: Índice de masa corporal, reconocimiento de patrones, características, procesamiento digital de imágenes.

Automatic calculation of body mass index using Artificial Vision

Abstract: Technological advances have been implemented in different fields of daily life; a field of great relevance is medicine. The investigation of methodologies that facilitate the early determination of symptoms and their subsequent treatment is of vital importance. This article proposes a method based on artificial vision for the calculation of body mass index (BMI), for this purpose people of different ages and complexions were gathered, from the images obtained from all people the characteristics of the area are obtained of interest to obtain a good BMI classification performance, obtaining a classification accuracy of 73%.

Keywords: Body mass index, pattern recognition, features, classification, digital image processing.

1. Introducción

Las nuevas tecnologías ayudan a las personas cada vez más a estar todo el tiempo conectadas (Mekki, Bajic, Chaxel & Meyer, 2019) (Laitinen & Valo, 2018), a poder movilizarse y a desarrollar sus actividades diarias (Poggi & Mattocchia, 2016) (Benevolo, Dameri & D'Auria, 2016) (Cervantes, García-Lamont, Rodríguez-Mazahua, Rendon, & Chau, 2016). Estamos incluyendo la tecnología en cada aspecto de nuestra vida llegando a un nivel de dependencia de estas. Dentro de toda la tecnología usada, la visión artificial abarca diversas líneas de investigación y aplicación, lo podemos ver en control de calidad de manufactura o maquinaria para automatizar cosechas (Suárez, López, Rivera & Ledezma, 2019) (Bejarano, Calvo & Henao, 2018) o diversas aplicaciones en la agricultura (Jalili, Morales, Cervantes, & Ruiz-Castilla, 2016) (Cervantes, Taltempa, Lamont, Castilla, Rendon & Jalili, 2017) (Cervantes, García-Lamont, Rodríguez-Mazahua, Rendon, & Chau, 2016) (Cervantes, Garcia Lamont, Rodriguez Mazahua, Zarco Hidalgo & Ruiz Castilla, 2018). Aplicaciones más usadas son los filtros en fotos digitales, pero también se realizan investigaciones para crear aplicaciones basadas en visión artificial que sirvan en el área de la salud, que es un área muy importante de estudio, donde se realizan investigaciones para poder detectar automáticamente ciertas enfermedades que se pueden observar en imágenes clínicas, resonancias magnéticas, entre otras (Abuzaghle, Barkana, & Faezipour, 2015).

La visión artificial tiene una gran importancia dentro de las nuevas tecnologías, ya que a partir de imágenes digitales se obtiene información necesaria para realizar ciertos procesos. A partir de la visión artificial se pueden automatizar procesos de cosechas, maquila de ropa, control de calidad de productos, construcción, etc. En el área de la salud la visión artificial puede ayudar a automatizar estudios que actualmente llevan mucho tiempo ya que tienen que ser interpretados por especialistas, o estudios que requieren equipos costosos o caros para realizarlos. Estudios como el cálculo del índice de masa corporal a partir de imágenes es muy importante debido a que la altura y el peso corporal son rasgos biométricos. Aunado a ello, estos pueden ser utilizados como un indicador de salud, pues regularmente el exceso de peso es asociado con diversas enfermedades como diabetes y enfermedades cardiovasculares. Un sistema automático para cálculo del IMC puede ser utilizado cuando no se tiene acceso a dispositivos de medición, ayudar al diagnóstico remoto de pacientes, evitar la necesidad de un diagnóstico médico en pacientes con problemas de sobrepeso o pacientes con problemas de bulimia (Wen & Guo, 2013) (Mardolkar, 2017) (Millard, Davies, Tilling, Gaunt & Smith, 2017). Sin embargo, la estimación del IMC de forma remota es un problema desafiante hoy en día. En este artículo se presenta un sistema de visión para la detección automática de IMC a partir de imágenes.

2. Estado del arte

Existen métodos rápidos para poder dar una primera impresión del estado de salud general de la persona, uno de ellos es el semblante que refleja la piel, el color de las escleróticas de los ojos, y también de manera muy importante la complexión corporal o índice de masa corporal (IMC).

El IMC se usa como factor determinante de la condición nutrimental de las personas, y se calcula a partir de la relación entre el peso y su estatura. Esta medida es usada para saber si la persona esta saludable o tiene algún problema de nutrición, pero también en base al IMC se puede determinar si la persona necesita estudios especializados de alguna posible enfermedad o solo asistencia nutricional. El índice de masa corporal (IMC) es la medida calculada a partir de la relación entre la altura y el peso de la persona, con el valor de IMC se puede clasificar al individuo dentro de una categoría como se muestra en la Tabla 1.

El IMC es una herramienta muy útil utilizada durante años como uno de los factores dentro del cuadro clínico de las personas, este aspecto lo vuelve un factor de estudio por si solo para poder obtener el cálculo de manera eficiente y pronta, no solo por especialistas de la salud, sino también para investigadores de áreas como las matemáticas (Ares,2013) (Mamat, Deraman, Mohamad Noor & Zulkifli, 2013) (Boutayeb, Boutayeb, Lamliili & Mohamed, 2014), especialistas del deporte (Duncan, Nevill, Woodfield & Al-Nakeeb, 2010) (Pontaga & Židens, 2011) (Douglas & Feresu 2013) (Mwangi & Rintaugu, 2017), ciencias de la computación (Wen & Guo, 2013) (Mardolkar,2017) (Millard, Davies, Tilling, Gaunt & Smith, 2017), que intentan crear una herramienta para automatizar la obtención del IMC. Actualmente se cuenta con herramientas web que calculan el IMC a partir de los parámetros solicitados.

Se han realizado investigaciones para perfeccionar el cálculo del IMC como en Bipembi (Bipembi, Acquah, Panford & Appiah, 2015), donde se propone un complemento al modelo matemático del cálculo para obtener mayor precisión de IMC, a partir de imágenes segmentadas realiza el cálculo, obteniendo un resultado con una diferencia de 4.04 con respecto a la manera clásica de cálculo. En sus experimentos utilizó solo 50 imágenes sin definir género de las personas. Otra investigación basada en cálculo de IMC a partir de procesamiento de imágenes es el realizado por Wen (Wen & Guo, 2013), su investigación se basa principalmente en el rostro del individuo, mediante un algoritmo de detección de rostros determina la posición de los ojos y determina la complejión del individuo a partir del tamaño del rostro. En sus experimentos dividen sus conjuntos de datos de acuerdo a la edad, género y grupo étnico. El conjunto de datos utilizado contiene 55,000 imágenes.

Otra investigación basada en el estudio de rostro para determinar el IMC del individuo es el realizado por Bolukbaş (Bolukbaş, Başaran & Kamaşak, 2019), en esta investigación implementa *Deep Learning* con una red neuronal para la clasificación, obteniendo una diferencia de 0.63 en la correlación de Pearson. Los autores utilizan una base de datos con 4206 imágenes de personas sin definir genero. En (Das, Dantcheva & Bremond, 2018) se utilizó un conjunto de datos con 1026 imágenes correspondientes a la misma cantidad de personas, la mitad de hombres y la mitad de mujeres, mediante un modelo Res-Net se clasificaron las imágenes obteniendo 0.60 en correlación de Pearson.

Una investigación similar usando rostros para el cálculo del IMC se realizó en Polania (Polania, Wang & Fung, 2018), donde se utilizó un repositorio de rostros para entrenar un modelo de redes siamesas para la clasificación del IMC, obteniendo una precisión de 0.496 y error absoluto medio de 0.554. El repositorio utilizado contiene 67756 imágenes de entrenamiento, 9045 de validación y 20,000 de prueba, sin especificar género.

Una investigación realizada por Shu (Shu, & Zhang, 2015) intenta determinar la salud de un individuo mediante imágenes a color del rostro de las personas, usando 3 conjuntos de datos (personas sanas, sobrepeso, diabetes), se utiliza la distancia euclidiana entre 4 regiones del rostro basándose en una gama de colores de 6 tonalidades, donde las personas sanas presentaban menor distancia entre el centroides de la tonalidad y el color de piel del individuo, obteniendo que la mayor variación de distancias se encontró en personas con sobrepeso.

Se han realizado investigaciones basadas en dispositivos de modelación 3d como el Kinect, en Tai (Tai and Lin, 2015) se utiliza el Kinect para obtener la representación facial de las personas, obteniendo una forma 2d del rostro y calcula las distancias desde los ojos al contorno del rostro y las curvaturas del mismo, concluyendo que la característica con mejor desempeño para la clasificación es el “sloperatio” es decir la distancia entre puntos adyacentes en el contorno del rostro.

Otra investigación donde se utiliza Kinect para la representación corporal es (Nahavandi, Abobakr, Haggag, Hossny, Nahavandi & Filippidis, 2017), en esta investigación se recrean figuras 3D con diferentes volúmenes para determinar su IMC, se basa en el cálculo de la superficie corporal usada por la representación, obteniendo una precisión de 95%. En sus experimentos generó 100,000 figuras 3D sintéticas.

En Madariaga (Madariaga & Linsangan, 2016) se implementa una metodología donde combinan el tratamiento de imágenes con dispositivos bluetooth, primero de la imagen se obtiene la altura de la persona, para calcular el peso se utiliza un dispositivo que crea un voltaje, donde se encuentra el cuerpo existe una variación, mediante una metodología independiente se toma ese valor y se calcula el peso del individuo, obteniendo diferencias entre el valor calculado y la estimación mediante el método propuesto es de menos del 3.50%. Sin embargo, no provee información sobre el tamaño del conjunto de datos utilizado.

En (Uyun & Efendi, 2019), se realiza el cálculo del peso ideal usando el IMC. Implementa una combinación de cuatro diferentes algoritmos y tratamiento de imágenes, en esta investigación se asume al cuerpo humano como una representación de tubo elíptico, entonces, la imagen sirve para determinar el área de la superficie ocupada por el supuesto tubo, en donde, previamente se calcula, en pixeles, la altura y el ancho del objeto dentro de la imagen. Se utilizó un conjunto de 47 imágenes con perspectiva frontal al objeto, capturadas desde la cámara de un teléfono inteligente Android, para clasificar el peso humano basado en cuatro clases, esas son; delgadas, ideales, gordas y obesas. De los cuatro algoritmos diseñados y probados, uno de ellos obtuvo una precisión de 78.7%.

Supranata (Supranata, Davin, Jeremy, Pratiwi & Wulandari, 2018), utiliza y compara el volumen del tubo elíptico y el área de superficie corporal para realizar el cálculo del peso utilizando imágenes frontales y laterales con un objeto humano. Para el volumen del tubo elíptico, el peso corporal del ser humano se mide aplicando la fórmula del volumen de un tubo, para el área de superficie corporal, el peso corporal del humano se mide aplicando el área de superficie de un tubo. el conjunto de 49 imágenes utilizado sirvió para obtener la precisión del método de volumen de tubo elíptico del 93,271% y la precisión del método del área de la superficie corporal del 92,854%. En ambas investigaciones se hace referencia a la obtención del peso del humano, sin embargo, no se clasifica con los valores propuestos por la OMS (Organización Mundial de la

Salud) para identificar y encasillar valores perjudiciales para la salud utilizando el IMC, con índices que van desde el infra-peso hasta la obesidad nivel III. En Supranata (Supranata, Davin, Jeremy, Pratiwi & Wulandari, 2018), se toma en cuenta el valor asociado a la imagen lateral para calcular el área de la superficie y el volumen de un tubo, no así en Uyun (Uyun & Efendi, 2019), que solo considera la posición frontal del humano para calcular el área de la superficie, perdiendo la posibilidad de tener una estrecha aproximación al diámetro del tubo que representar el objeto en la imagen. Una investigación reciente (Amador, Cervantes, Jalili, & Castilla, 2018) muestra un sistema similar al propuesto en este artículo, sin embargo, solo se utiliza una imagen por cada persona. Creemos que agregando imágenes laterales y traseras al conjunto de datos la precisión mejorará.

3. Metodología propuesta

En esta sección se detalla la manera en que se recopila la información necesaria para el cálculo automático de IMC como se muestra en la Figura 1. Inicialmente se obtienen las imágenes de las personas, posteriormente se realiza un pre-procesamiento a las imágenes obtenidas se realiza la segmentación del área de interés, se realiza la extracción de características para obtener un clasificador eficaz para el cálculo del IMC.

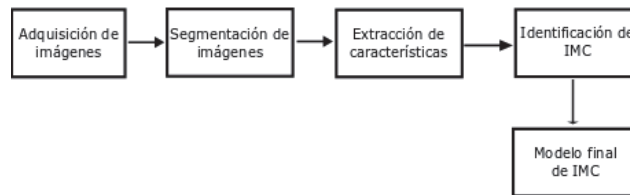


Figura 1 – Metodología propuesta

3.1. Adquisición de Imágenes

Para esta investigación se reunió una diversidad de personas de diferentes edades y peso. Se solicitó que las personas vistieran ropa oscura, se colocó una tela color blanco de fondo. La cámara utilizada fue una cámara digital Nikon Coolpix P510(16.1MP).

A cada individuo se le tomo el peso utilizando un analizador de composición corporal (Mod. 230), su estatura y edad, posteriormente descalzo se le tomaron 3 imágenes, una imagen de frente, una de perfil y una de espaldas, en total se fotografió a un total de 122 personas y se obtuvieron 366 imágenes en formato JPG de $N \times N$ píxeles. Como se muestra en la Figura 3. Como se muestra en la Tabla 1, la distribución de los individuos es homogénea en todas las categorías, siendo las últimas tres categorías la que menos ejemplos tienen debido a la falta de personas con esas características.

3.2. Segmentación de imágenes

Para el estudio de las imágenes se puede dividir a estas en dos clases principales, el área de interés de las imágenes y el fondo. Para poder seleccionar el área de interés

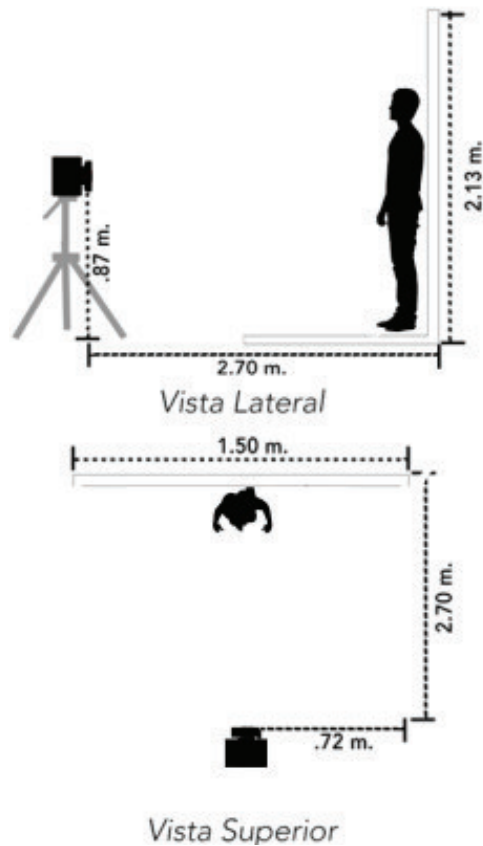


Figura 2 – Disposición de la cámara con respecto al individuo

de las imágenes existen diversas técnicas utilizadas, las cuales principalmente están basadas en imágenes en formato RGB y escala de grises. En el método propuesto una vez realizado el pre-procesamiento de las imágenes, estas se pasaron a escala de grises y se implementó el método de Otsu (Otsu,1979), ya que en la literatura este método es utilizado obteniendo muy buenos resultados en la segmentación de imágenes. En el área de visión artificial existen técnicas que ayudan a mejorar la calidad de las imágenes que se están tratando, es decir, técnicas que ayudan a aumentar el brillo de la imagen, eliminar el emborronamiento, aumentar o disminuir el tamaño, entre otros. En esta investigación a las imágenes obtenidas se les aplicaron filtros promedio, el cual suaviza la imagen y filtro gaussiano que permite eliminar ruido en la imagen. Posteriormente se implementó segmentación por el método de Otsu para obtener la región de interés, que para esta investigación es la silueta corporal. En la Figura 4. Se ejemplifica el proceso de tratamiento de imágenes tomada a los individuos.

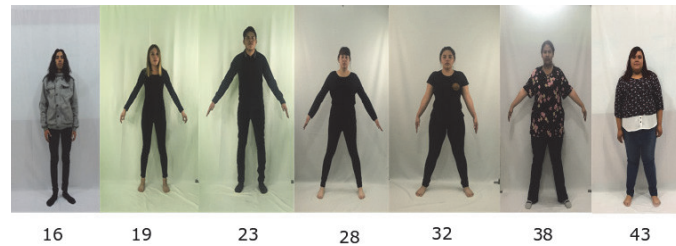


Figura 4a – Imagen original de rangos de IMC



Figura 4b – Imagen umbralizada

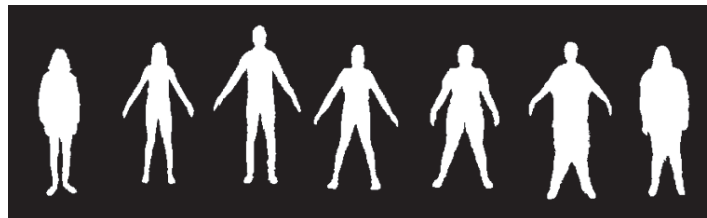


Figura 4c – Silueta final

3.3. Extracción de características

En el tratamiento de imágenes se utilizan principalmente las características geométricas. Las características geométricas son una de las más importantes para la clasificación de objetos. Las características geométricas proveen información de la forma y dimensiones de la región de interés de la imagen (Zhang, 2011) (Hu, 1962) (Hu & Collomosse, 2013) (Cervantes, Taltempa, Lamont, Castilla, Rendon & Jalili, 2017).

Con estas características se puede obtener el área, el perímetro de la región, su redondez, altura, ancho, rectangularidad, proyección (en sus componentes x e y), excentricidad, centro de gravedad (componentes X e y), orientación, solidez, extensión, área convexa, elipsidad (varianza, orientación, excentricidad, eje mayor, eje menor, centro de elipse). Un sistema eficiente de clasificación debería reconocer objetos sin importar la orientación, localización y tamaño, siendo invariantes al escalamiento, rotación y posición del objeto.

Los momentos son usados en el reconocimiento de imagen ya que son invariantes a la rotación y al tamaño del objeto. Los momentos fueron introducidos por Hu (Hu,1962). Los momentos de Hu de orden $(p + q)$ de una función de intensidad $f(x, y)$ se definen:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (1)$$

donde $f(x, y)$ representa la intensidad del pixel en la región segmentada. Los momentos de orden bajo describen la forma de la región. Por ejemplo, m_{00} describe el área de la región segmentada, mientras que 01 y 10 describen las coordenadas X e y del centro de gravedad. Sin embargo, los momentos m_{02} , m_{03} , m_{11} , m_{12} , m_{20} , m_{21} , m_{30} son invariantes a la translación, rotación e inversión. Los momentos centrales son invariantes al desplazamiento y se pueden calcular de la siguiente manera:

$$\mu_{pq} = \sum_{i,j \in R} (i - \bar{i})^p (j - \bar{j})^q \quad (2)$$

donde p, q describen el grado del momento, i, j definen las coordenadas de cada uno de los puntos de la región segmentada, mientras \bar{i} y \bar{j} representan los centros de gravedad o la media de los puntos que pertenecen a la región segmentada en los ejes coordenados x e y , estos también pueden ser obtenidos mediante:

$$\bar{i} = \frac{m_{10}}{m_{00}}, \bar{j} = \frac{m_{01}}{m_{00}} \quad (3)$$

Los momentos de Hu se pueden obtener de la siguiente manera:

$$\begin{aligned} \Phi_1 &= \eta_{20} + \eta_{02} \\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \Phi_4 &= (\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2 \\ \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + 3\eta_{03})^2] \\ &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 &= (3\eta_{21} - \eta_{03}) + (\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (4)$$

donde:

$$\eta_{pq} = \frac{\mu_{rs}}{\mu_{00}^t}, t = \frac{p+q}{2} + 1 \quad (5)$$

Algunas veces es necesario tener características invariantes a traslación, rotación, escala y transformaciones afines. Los momentos de Flusser obtienen características derivadas a partir de los momentos centrales de segundo y tercer orden que son invariantes a transformaciones afines (Flusser & Suk, 1993) (Cervantes, Taltempa, Lamont, Castilla, Rendon & Jalili, 2017), estas pueden ser obtenidas a partir de las siguientes ecuaciones:

$$I_1 = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4} \quad (6)$$

$$I_2 = \frac{\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2}{\mu_{00}^{10}} \quad (7)$$

$$I_3 = \frac{\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)}{\mu_{00}^7} \quad (8)$$

$$I_4 = (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 +$$

$$12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} -$$

$$8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21} + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} -$$

$$6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11} \quad (9)$$

Otras características usadas son descriptores de elipse, convexidad de la región, momentos de Flusser (F_1, \dots, F_4) (Flusser & Suk, 1993) (Mingqiang, Kidiyo & Ronsin, 2008), momentos R (R_1, \dots, R_{10}), descriptores de Fourier (Mingqiang, Kidiyo & Ronsin, 2008) (Nixon & Aguado, 2002)

En esta investigación se obtuvieron 57 características geométricas de la silueta de cada individuo. El vector de características resultante Xg puede ser representado de la siguiente manera:

$$Xg = [X_1, X_2, \dots, X_n] \quad (10)$$

$$Xg = [X_{gb}, X_{Hu}, X_F, X_R, X_{DF}] \quad (11)$$

donde X_{gb} representa las características geométricas elementales ($X_{gb} = X_1, \dots, X_{28}$), X_{Hu} representa los momentos de Hu ($X_{Hu} = X_{29}, \dots, X_{35}$), X_F representa los momentos invariantes de Flusser ($X_F = X_{36}, \dots, X_{39}$), X_R representa los momentos invariantes de cambios de iluminación ($X_R = X_{40}, \dots, X_{49}$), X_{DF} representa los primeros 8 descriptores de Fourier ($X_{DF} = X_{50}, \dots, X_{57}$).

3.4. Identificación de IMC

Para poder identificar a que rango pertenece el IMC se necesita reconocer los patrones específicos de cada conjunto de datos que los diferencian. El reconocimiento de patrones es una tarea muy compleja. Para la experimentación en esta investigación se utilizaron los siguientes clasificadores para poder comparar los resultados obtenidos.

3.4.1. SVM

Las SVM (Vapnik, 1995). son una de las técnicas de clasificación más utilizadas en los últimos años, Las características clave de las SVMs son el uso de *kernels* al trabajar en conjuntos no-lineales, la ausencia de los mínimos locales, la solución depende de un pequeño subconjunto de datos y el poder discriminativo del modelo obtenido al optimizar el margen de separabilidad entre clases, estas características permiten a las SVM obtener resultados muy competitivos en comparación con otros clasificadores. Formalmente las SVM pueden ser definidas de la siguiente manera:

Asumiendo que un conjunto de datos de entrenamiento X es dado como:

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \text{ i.e. } X = \{x_i, y_i\}_{i=1}^n \quad (12)$$

donde $x_i \in \mathcal{R}^d$ y $y_i \in \{1, -1\}$. Entrenar una SVM permite resolver un problema de programación cuadrática como sigue:

$$\max_{\alpha_i} -\frac{1}{2} \sum_{i,j=1}^l \alpha_i y_i \alpha_j y_j K(x_i \cdot x_j) + \sum_{i=1}^l \alpha_i \quad (13)$$

Sujeto a: $\sum_{i=1}^l \alpha_i y_i = 0$, $C \geq \alpha_i \geq 0$, $i = 1, 2, \dots, l$

Donde $C > 0$, $\alpha_i = [\alpha_i, \alpha_i, \dots, \alpha_i]^T$, $\alpha_i \geq 0$, $i = 1, 2, \dots, l$ son coeficientes que corresponden a x_i , x_i con α_i diferentes a cero, que son llamados Vectores Soporte(SV). La función K es una función. Las SVM son clasificadores lineales, es decir, se clasifica entre dos conjuntos de datos mediante la construcción de una línea que separa dos clases. Cuando esto no es posible, se utiliza una función llamada *kernel* que transforma el espacio de entrada a un espacio altamente dimensional, en donde los conjuntos pueden ser separados linealmente después de la transformación. Sin embargo, la elección de una función está restringida a aquellas que satisfacen las condiciones de Mercer (Vapnik, 1995).

Sea S el conjunto de SV obtenidos después del entrenamiento, entonces el hiperplano óptimo es dado por:

$$\sum_{i \in S} (\alpha_i y_i) K(\mathbf{x}_i \cdot \mathbf{x}_j) + b = 0 \quad (14)$$

Y la función de decisión optima es definida como

$$f(x) = \text{sign} \left(\sum_{i \in S} (\alpha_i y_i) (\mathbf{x}_i \cdot \mathbf{x}_j) + b \right) \quad (15)$$

Donde $\mathbf{x} = [x_1, x_2, \dots, x_i]$ son los datos de entrada, α_i y y_i son los multiplicadores de Lagrange. Un nuevo objeto \mathbf{x} puede ser clasificado empleando Eq(14). El vector \mathbf{x}_i es dado en la forma de producto punto. Existe un multiplicador de Lagrange α para cada punto de entrenamiento. Cuando el máximo margen del hiperplano es encontrado, solamente los puntos más cercanos al hiperplano satisfacen $\alpha > 0$. Estos puntos son los vectores

soporte (SV).

3.4.2. Redes neuronales

Una red neuronal es un conjunto de neuronas artificiales interconectadas que utilizan modelos matemáticos para procesar información (Werbos, 1994). Las múltiples conexiones entre las neuronas forman un sistema adaptivo cuyos pesos se actualizan mediante un particular algoritmo de aprendizaje. Las redes neuronales han sido utilizadas en numerosos campos de aplicación con distintos algoritmos de aprendizaje.

De los distintos algoritmos de aprendizaje, uno de los más utilizados es el algoritmo de aprendizaje *backpropagation* (BP) (Rumelhart, Hinton & Williams, 1986). Para llevar a cabo el proceso de aprendizaje, el algoritmo de aprendizaje BP cambia iterativamente los pesos entre las neuronas minimizando el error cuadrático entre la salida deseada y la obtenida con los pesos actuales. Cada uno de los ejemplos del conjunto de entrenamiento $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ son utilizados para ajustar los pesos en la red. Al ser presentado un ejemplo, la señal es propagada hacia adelante de la red hasta que la salida es obtenida. La salida de la j-ésima unidad oculta es calculada como:

$$O_{nj}^h = f_j^h(\text{net}_{nj}^h) = \frac{1}{1 + \exp(-\text{net}_{nj}^h)} \quad (16)$$

donde $\text{net}_{nj}^h = \sum w_{ji}^h x_{ni} + \theta_j^h$ representa el peso de la conexión de la i-ésima neurona de entrada a la j-ésima neurona oculta. θ_j^h y f_j^h representan el sesgo y la función de activación de la j-ésima neurona oculta. Por otro lado, la salida de la k-ésima neurona está dado por

$$O_{nk}^o = f_k^o(\text{net}_{nk}^o) = \frac{1}{1 + \exp(-\text{net}_{nk}^o)} \quad (17)$$

donde los superíndices h y o se refieren a las cantidades en las capas ocultas y de salida respectivamente. El error entre la salida actual y la salida deseada es calculado para ajustar los pesos mediante $E_n = \frac{1}{2} \sum_{k=1}^C (t_{nk} - o_{nk})^2$. El procedimiento de ajuste es obtenido a partir del método de gradiente descendente para reducir la magnitud del error. El procedimiento es primeramente aplicado a los pesos en la capa de salida y retropropagado a través de la red hasta que los pesos en la primera capa han sido ajustados $\Delta w_{kj}^o = -\alpha \frac{\partial E_n}{\partial w_{kj}^o}$ y $w_{ji}^h = -\alpha \frac{\partial E_n}{\partial w_{ji}^h}$. Este procedimiento es realizado para cada ejemplo en el conjunto de datos hasta que se cumple un criterio de paro.

3.4.3 Naive Bayes

Los clasificadores Bayesianos (Bernardo & Smith, 2000) (Russell & Norvig, 2002) se basan en la teoría de decisión de Bayes, El principio de Bayes provee una metodología fundamental para resolver problemas de clasificación de patrones cuando la distribución de probabilidad de los patrones es conocida. Un clasificador Bayesiano utiliza un enfoque probabilístico para asignar la clase a un ejemplo. Sea C la clase de un objeto que pertenece a un conjunto de m clases (C_1, C_2, \dots, C_m) y X_k un objeto con k características $X_k = [x_1, x_2, \dots, x_k]$, en nuestro caso es el conjunto de características que definen a una hoja. el clasificador Bayesiano, calcula la probabilidad condicional posterior $P(C_i | X_k)$ utilizando la regla de Bayes:

$$P(C_i | X_k) = \frac{P(X_k | C_i) P(C_i)}{P(X_k)}, \quad i = 1, 2, \dots, m. \quad (18)$$

En la ecuación Eq(11) $P(X_k | C_i)$, $P(C_i)$, y $P(X_k)$ son calculadas a partir de los datos de entrenamiento. De acuerdo a la teoría de Bayes para una observación dada X_k , la clase a la que pertenece está dada por la probabilidad posterior máxima:

$$f(X_k) = \arg \max_i P(C_i | X_k) \quad (19)$$

3.4.4 Regresión logística

La regresión logística (Ng, & Jordan, 2002) es utilizada para modelar las probabilidades de clase posterior sin tener que aprender las densidades de clase condicional, facilitando la clasificación en pequeños conjuntos de entrenamiento y una menor complejidad. Sea $\pi_i = p(Y_i = 1 | X_i)$ donde X_i es un vector de tamaño $1 \times (p + 1)$ con el primer elemento igual a 1 y los elementos restantes corresponden a las características extraídas de la silueta para el ejemplo i . El modelo de regresión logística relaciona π_i con las características mediante la función:

$$\text{logit}(\pi_i) = \log \frac{\pi_i}{1 - \pi_i} = X_i \beta \quad (20)$$

donde $\beta = (\beta_0, \beta_2, \dots, \beta_p)$ es el vector de coeficientes de regresión.

4. Resultados experimentales

En esta sección se desarrolla el proceso de experimentación del cálculo del IMC para la creación de un clasificador de IMC. Para la experimentación se normalizaron los datos, se utilizó validación cruzada (*cross validation*) con $k=10$, se utilizaron datos normalizados. El desempeño de los clasificadores se llevó a cabo utilizando Weka. Los resultados obtenidos con los clasificadores utilizados se muestran en la Tabla 2 donde se visualiza el AUC, F-measure, y Acc de cada uno de los clasificadores.

IMC	Altura promedio(cm)	Peso promedio(KG)	Edad promedio	Hombres	Mujeres	Numero de muestras
17-19	164	49	22	7	11	18
20-21	166	56	20	10	10	20
22-23	165	61	22	16	12	28
24-25	160	62	21	5	19	24
26-27	165	70	22	8	11	19
28-29	158	69	30	1	4	5
30-31	167	86	27	2	2	4
32-33	170	88	21	2	2	4

Tabla 1 – Distribución y promedio de individuos

Los resultados presentados en esta Tabla muestran el desempeño obtenido solo con imágenes de frente. Como se Puede observar en los resultados anteriores, los cuatro clasificadores tuvieron resultados similares en su desempeño, sin embargo, el que mejor desempeño mostro sobre todos fue el clasificador SVM con un 85.1% de clasificación, y el que menor desempeño mostro fue Regresión Lineal con 78%. El rango de IMC que menor precisión obtuvo fue el de 32-33 y el que mejor resultados tuvo fue el rango 26-27. Esto es entendible por el número de imágenes utilizadas para cada clase.

Categoría	Bayes			SVM _{RBF}			BP			RL		
	Acc	F-measure	AUC	Acc	F-measure	AUC	Acc	F-measure	AUC	Acc	F-measure	AUC
17-19	87.13	0.87	0.86	89.47	0.893	0.873	86.2	0.83	0.85	85.3	0.82	0.85
20-21	75.9	0.83	0.89	77.34	0.838	0.920	76.6	0.82	0.87	76.7	0.79	0.83
22-23	71.8	0.77	0.62	75.1	0.821	0.671	72.7	0.78	0.58	73.5	0.78	0.59
24-25	77.3	0.80	0.91	79.4	0.835	0.947	78.5	0.81	0.89	77.9	0.80	0.88
26-27	82.7	0.72	0.90	85.7	0.857	0.924	83.2	0.79	0.87	81.7	0.78	0.81
28-29	64.7	0.75	0.83	65.7	0.789	0.835	65.3	0.76	0.81	61.3	0.71	0.79
30-31	55.9	0.73	0.79	58.4	0.735	0.826	56.3	0.71	0.79	51.9	0.67	0.78
32-33	43.8	0.51	0.74	46.2	0.610	0.815	42.7	0.56	0.75	40.3	0.51	0.73
Promedio	69.90	0.7475	0.81	72.16	0.7972	0.851	70.18	0.7575	0.80	68.57	0.7325	0.78

Tabla 2 – Desempeño de clasificadores con imágenes de frente

Categoría	Bayes			SVM _{RF}			BP			RL		
	Acc	F-measure	AUC	Acc	F-measure	AUC	Acc	F-measure	AUC	Acc	F-measure	AUC
17-19	86.4	0.89	0.89	90.2	0.92	0.91	86.5	0.82	0.84	87.2	0.84	0.89
20-21	78.2	0.85	0.92	78.3	0.85	0.93	77.3	0.78	0.85	77.2	0.82	0.87
22-23	74.6	0.80	0.74	76.6	0.85	0.72	73.2	0.75	0.62	74.5	0.81	0.86
24-25	78.1	0.81	0.94	80.1	0.86	0.95	77.9	0.83	0.92	78.2	0.83	0.92
26-27	85.1	0.73	0.90	86.8	0.87	0.97	81.4	0.83	0.91	82.4	0.80	0.84
28-29	63.9	0.77	0.88	67.2	0.79	0.85	66.5	0.68	0.79	63.7	0.72	0.81
30-31	58.4	0.78	0.85	54.2	0.72	0.87	60.4	0.59	0.76	54.6	0.69	0.79
32-33	46.4	0.54	0.76	52.5	0.67	0.82	45.9	0.48	0.70	42.6	0.54	0.78
Promedio	71.38	0.7712	0.86	73.23	0.8162	0.8775	71.13	0.72	0.7987	70.0	0.756	0.845

Tabla 3 – Desempeño de clasificadores con imágenes de frente, lateral y atrás

La Tabla 3. Muestra los resultados con las 3 imágenes por persona, es claro ver como la precisión mejora para casi todos los clasificadores. Solo con el algoritmo Backpropagation la precisión cae un poco con respecto a los resultados con una sola imagen (Ver Tabla 2). Sin embargo, podría deberse a una mala selección de parámetros. Aunque la mejora es clara, esta no es significativa. Esto podría deberse al número de imágenes, así como el desbalance en los conjuntos de datos. En este artículo trabajamos con un conjunto de datos muy pequeño y no balanceado. Sin embargo, como trabajo futuro sería importante trabajar con un conjunto de datos balanceado y con más imágenes. Esto mejoraría significativamente el desempeño de los clasificadores.

5. Conclusiones

En este se propone un método para calcular el IMC de los individuos mediante las características obtenidas de las imágenes de la silueta de las personas. De esta manera se obtiene el IMC sin incomodar ni afectar a las personas que sufren por su peso. Los resultados obtenidos en relación con investigaciones similares son muy buenos, sin embargo, los resultados por categoría muestran que los datos de ejemplos utilizados son insuficientes ya que la media de las personas no presenta ese grado de obesidad y es difícil de conseguir imágenes de personas con IMC alto. Otro punto importante observado en los experimentos realizados es que el IMC depende del genero de la persona. Es muy probable que dividiendo la clasificación por genero la precisión del sistema de visión propuesto mejore su desempeño. En los experimentos presentados en este artículo es posible ver que al aumentar el número de imágenes de cada persona (imágenes lateral y trasera) la precisión es mejorada. Sin embargo, la mejora no es significativa y la principal desventaja es que el tiempo de procesamiento aumenta enormemente.

En trabajos futuros se pretende mejorar la investigación en tres aspectos:

1. Base de imágenes. Es posible mejorar el desempeño del sistema aumentando el número de personas en la base de imágenes. En los experimentos se puede inferir además que dividiendo el proceso de clasificación del IMC por genero se puede

mejorar aún más el desempeño de los clasificadores. Una desventaja mayor es que las clases utilizadas están desbalanceadas. Mientras que de algunos rangos se tienen muchas imágenes, de otros rangos se tienen muy pocas imágenes. Para trabajo futuro se debe homogeneizar el número de personas en cada rango.

2. La segmentación. En los experimentos realizados fue necesario eliminar bastantes imágenes de la base de datos original. Esto debido a que al segmentar la imagen no se obtenía perfectamente el contorno de la imagen. Cualquier brillo en la imagen puede provocar una mal segmentación u obtención del contorno de la imagen, disminuyendo la precisión de identificación. En los trabajos futuros se pretende mejorar un método de segmentación que enfrente los retos anteriores.
3. Selección de características. En los experimentos realizados se obtuvieron las características definidas en la Sección 3.2. sin embargo, es posible que algunas de estas sean innecesarias o incluso peor que introduzcan ruido en el clasificador. Como trabajo futuro se utilizarán técnicas para seleccionar adecuadamente las mejores características, eliminando aquellas innecesarias y las que introduzcan ruido. Esto podría mejorar la precisión y disminuir los tiempos.

Referencias

- Abuzaghle, O., Barkana B.D., & Faezipour M. (2015). Noninvasive Real-Time Automated Skin Lesion Analysis System for Melanoma Early Detection and Prevention. *IEEE Journal of Translational Engineering in Health and Medicine* 3, 1–12.
- Amador, J. D., Cabrera, J. E., Cervantes, J., Jalili, L. D., & Castilla, J. S. R. (2018, October). Automatic Calculation of Body Mass Index Using Digital Image Processing. In *Workshop on Engineering Applications* (pp. 309-319). Springer, Cham.
- Ares, G. (2013). *Mathematical and statistical methods in food science and technology*, Willey, J. And sons (2013)
- Bejarano, M., Calvo A.F. & Henao C.A. (2016). Supervised learning models for control quality by using color descriptors: A study case, in: 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA). Presented at the 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA), pp. 1–7.
- Benevolo, C., Dameri R.P. & D'Auria B., (2016). Smart Mobility in Smart City, in: Torre, T., Braccini, A.M., Spinelli, R. (Eds.), *Empowering Organizations, Lecture Notes in Information Systems and Organisation*. Springer International Publishing, Cham, pp. 13–28.
- Bernardo, S. and Adrian F M Smith (2000). *Bayesian Theory*. John Wiley & Sons.
- Bipembi, H. a.-A, J. B. Hayfron-Acquah, Joseph K. Panford & Obed Appiah (2015). Calculation of Body Mass Index using Image Processing Techniques. *International Journal of Artificial Intelligence and Mechatronics*, 4(1), 1-7.
- Bolukbaş, G., Başaran E. & Kamaşak M. (2019). BMI Prediction From Face Images, in: 2019 27th Signal Processing and Communications Applications Conference (SIU). pp. 1–4.

- Boutayeb, W., Boutayeb A., Lamlili M., D. Mohamed (2014). A mathematical model of overweight/obesity in Morocco, using human biomass. *International Journal of Latest Research on Science and Technology*, 3(6), 65-67.
- Cervantes, J., García-Lamont, F., Rodríguez-Mazahua, L., Rendon, A.Y., Chau, A.L. (2016) Recognition of Mexican sign language from frames in video sequences. In: Huang, D.-S., Jo, K.-H. (eds.) ICIC 2016. LNCS, vol. 9772, pp. 353–362. Springer, Cham (2016).
- Cervantes, J., Taltempa, J., Lamont, F. G., Castilla, J. S. R., Rendon, A. Y., & Jalili, L. D. (2017). Análisis comparativo de las técnicas utilizadas en un sistema de reconocimiento de hojas de planta. *Revista Iberoamericana de Automática e Informática industrial*, 14(1), 104-114.
- Cervantes, J., Garcia Lamont, F., Rodriguez Mazahua, L., Zarco Hidalgo, A., Ruiz Castilla, J.S. (2018). Complex identification of plants from leaves. In: Huang, D.-S., Gromiha, M.M., Han, K., Hussain, A. (eds.) ICIC 2018. LNCS (LNAI), vol. 10956, pp. 376–387. Springer, Cham (2018).
- Chih-Hua Tai and Daw-Tung Lin (2015). A Framework for Healthcare Everywhere: BMI Prediction Using Kinect and Data Mining Techniques on Mobiles, in: 2015 16th IEEE International Conference on Mobile Data Management. pp. 126–129.
- Das A., Dantcheva A., & Bremond F. (2018), Mitigating Bias in Gender, Age and Ethnicity Classification: a Multi-Task Convolution Neural Network Approach. The European Conference on Computer Vision (ECCV) Workshops.
- Douglas, D. & Shingairai A Feresu (2013). The relationship between physical activity, body mass index, and academic performance and college-age students. *Open Journal of Epidemiology*. 3, 4-11.
- Duncan, M.J., Alan Nevill, Lorayne Woodfield & Yahya Al-Nakeeb (2010). The relationship between pedometer-determined physical activity, body mass index and lean body mass index in children *International Journal of Pediatric Obesity*, 5: 445–450.
- Flusser, J. & Tomás Suk (1993). Pattern recognition by affine moment invariants. *Pattern Recognition* 26(1), 167--174.
- Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, 8(2), 179-187.
- Hu, R. & John Collomosse (2013). A performance evaluation of gradient field HOG descriptor for sketch based image retrieval. *Computer Vision and Image Understanding*. (2013) 790--806.
- Jalili, L.D., Morales, A., Cervantes, J., Ruiz-Castilla, J.S. (2016). Improving the performance of leaves identification by features selection with genetic algorithms. In: Figueroa-García, J.C., López-Santana, E.R., Ferro-Escobar, R. (eds.) WEA 2016. CCIS, vol. 657, pp. 103–114. Springer, Cham (2016).

- Laitinen, K. & Valo M., (2018). Meanings of communication technology in virtual team meetings: Framing technology-related interaction. *International Journal of Human-Computer Studies* 111, 12–22
- Madariaga, N. & Noel B. Linsangan (2016). Application of artificial Neural Network and Background Subtraction for determining BMI in Android Devices Using Bluetooth
- Mamat, M., Siti Kamila Deraman, Noor Maizura Mohamad Noor and Nor Fadzilah Zulkifli (2013). Relationship between Body Mass Index and Healthy Food with a Balanced Diet, *Applied Mathematical Sciences*, Vol. 7, 2013, no. 4, 153 – 159
- Mardolkar, M. (2017). Body Mass Index (BMI) Data Analysis and Classification. *International Journal of Computer Science and Mobile Computing*, 6(2), 8-16.
- Mekki, K., Bajic E, Chaxel F & Meyer F. (2019). A comparative study of LPWAN technologies for large-scale IoT deployment. *ICT Express* 5, 1–7.
- Millard, L. Davies NM, Tilling K, Gaunt TR and Davey Smith G (2017). Searching for the causal effects of BMI in over 300 000 individuals, using Mendelian randomization. *bioRxiv preprint first posted online Dec. 19, 2017.*
- Mingqiang, Y., Kpalma Kidiyo and Ronsin Joseph (2008). *A Survey of Shape Feature Extraction Techniques, Pattern Recognition Techniques*, INTECH Open Access Publisher.
- Mwangi, F. & Elijah G. Rintaugu (2017). Physical Activity and Health Related Physical Fitness Attributes of Staff Members in a Kenyan Public University. *International Journal of Sports Science*, 7, 81-86.
- Nahavandi, D., A. Abobakr, H. Haggag, M. Hossny, S. Nahavandi & D. Filippidis (2017). A skeleton-free kinect system for body mass index assessment using deep neural networks. 1-6. [10.1109/SysEng.2017.8088252](https://doi.org/10.1109/SysEng.2017.8088252).
- Ng, A. & Michael I. Jordan (2002). On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes. *Adv. Neural Inf. Process. Sys*, 841-848.
- Nixon, M. & Alberto S. Aguado (2002). *Feature Extraction and Image Processing*, Academic Press.
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histogram. *IEEE Trans. Systems Man, and Cybernetics*, Vol. 9, pp. 62-66.
- Poggi, M. & Stefano Mattocchia (2016). A wearable mobility aid for the visually impaired based on embedded 3D vision and deep learning, in: *2016 IEEE Symposium on Computers and Communication (ISCC)*. Presented at the 2016 IEEE Symposium on Computers and Communication (ISCC), pp. 208–213.
- Polania, L., Dongning Wang, Glenn Fung (2018). Ordinal Regression Using Noisy Pairwise Comparisons for Body Mass Index Range Estimation. (WACV). pp. 782–790.

- Pontaga, I. & Jānis Žīdens (2011). Estimation of Body Mass Index in Team Sports Athletes. *Lase Journal of Sport Science*, 2(2), 33-44.
- Rumelhart D.E., Geoffrey E. Hinton & Ronald J. Williams (1986). Learning representations by back-propagating errors, *Nature*, 323(6088): 533--536.
- Shofwatul 'Uyun & Toni Efendi (2019) Classification of Human Weight Based on Image. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)* Vol. 13, No.2 April 2019, pp. 105~116, DOI: 10.22146/ijccs.35794
- Supranata, T. H., Davin, P. S. S., Jeremy, D. K., Pratiwi, A. E., & Wulandari, M. (2018). Body Weight Measurement Using Image Processing Based on Body Surface Area and Elliptical Tube Volume. In 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE) (pp. 290-294). IEEE.
- Stuart Russell & Peter Norvig (2002). *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- Suárez, L.J., López Y.P., Rivera W.F., Ledezma A. (2019). Silkworm Growth Monitoring in Second Stage -Instar- Using Artificial Vision Techniques, in: Corrales, J.C., Angelov, P., Iglesias, J.A. (Eds.), *Advances in Information and Communication Technologies for Adapting Agriculture to Climate Change II*, *Advances in Intelligent Systems and Computing*. Springer International Publishing, Cham, pp. 58–72
- Ting Shu, & Bob Zhang.(2015). Facial color analysis of Overweight-Obesity and its relationship to Healthy and Diabetes Mellitus using statistical pattern recognition. *TENCON 2015*pp. 1–4
- Vapnik,V.(1995). *The Nature of Statistical Learning Theory*. Springer.
- Wen,L. & Guo G.,(2013). A computational approach to body mass index prediction from face images. *Image and Vision Computing* 31 (2013) 392–400
- Werbos, P.J.(1994). *The Roots of Backpropagation. From Ordered Derivatives to Neural Networks and Political Forecasting*. New York, NY: John Wiley & Sons, Inc.
- Zhang, S. (2011). Modified locally linear discriminant embedding for plant leaf recognition, *Neurocomputing*, 74(14), 2284-2290.