

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
FACULTAD DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA,
INFORMÁTICA Y MECÁNICA
ESCUELA PROFESIONAL DE INGENIERÍA INFORMÁTICA Y DE SISTEMAS



TESIS

**“ANÁLISIS DE MÉTODOS DE VISIÓN COMPUTACIONAL Y
MACHINE LEARNING PARA LA CLASIFICACIÓN DE
IMÁGENES DE VARIEDADES DE PAPA NATIVA”**

Para optar al título profesional de:
INGENIERO INFORMÁTICO Y DE SISTEMAS

Presentado por:
Br. YENNIFER ZULMA CALVO CCALLO
Br. CARLA DORIS CARDOSO CUSIHUALLPA

Asesor:
M.Sc. LAURO ENCISO RODAS

Co-asesora:
M.Eng. EFRAINA GLADYS CUTIPA ARAPA

CUSCO - PERÚ
2018

RESUMEN

La poca difusión de las diversas variedades de papas nativas existentes en el mercado repercuten en el consumo limitado. Si bien es cierto que, existen diversas investigaciones sobre este tubérculo en distintas áreas como: la Agricultura, Genética, Salud, entre otros. El área de las Ciencias de la Computación también forman parte de ello, específicamente en el área de Visión Computacional, que ha tomado importancia en el control de calidad y/o enfermedades de la papa, mas no investigaciones para clasificar automáticamente papas nativas, ni investigaciones sobre cuales son los métodos más apropiados de visión computacional y *machine learning* para el mismo fin, por ser un producto propio de los Andes del Perú.

En el desarrollo del proyecto se analiza los métodos de visión computacional y *Machine Learning* para clasificar variedades de papas nativas. Para la clasificación manual de las variedades de papas nativas, se utiliza características externas de color, forma y textura. Se propone un extractor que contempla dichas características. Se obtiene características de color en los espacios: RGB, CMY, HSV, HSI, YCbCr, YIQ, LUV y Colores Oponentes. Se obtiene características de forma con los descriptores de Fourier Polar Bidimensionales, por ser invariantes a la rotación, traslación y escalamiento, sin perder la información de la imagen. Se obtiene características de textura realizando un análisis estadístico utilizando la matriz de coocurrencia, por ser el método más popular para extraer información importante de la textura. También se estudia los métodos de extracción de características *Scale Invariant Feature Transform* y *Speed Up Robust Features*; proponiéndose mejorar su rendimiento, al incluir el color en sus descriptores. Dentro de los métodos de *Machine Learning* se estudia los clasificadores *Support Vector Machine* y *Random Forest*.

Se demuestra que al incluir la característica de color en los métodos *Scale Invariant Feature Transform* y *Speed Up Robust Features* muestran una mejor precisión que los métodos sin la característica de color. El método con mejor tasa de aciertos para la clasificación de papa nativa es el extractor propuesto con el clasificador *Random Forest*, demostrando que el extractor propuesto es apropiado para clasificar las variedades de papa nativa.

Palabras claves: Clasificación de papa nativa, espacios de colores, Matrix de Coocurrencia, Descriptores de Fourier, SIFT, SURF, SVM, *Random Forest*.