An image can be mathematically represented by a function that takes in a matrix of red, green, blue, and greyscale values and converts them into pixels in some fixed display (Frery et al, 2) . Therefore an image (a visual representation of a matrix) can be represented in the following way (where S represents total dimension of the matrix (or coordinate system of the fixed display) f : S → K, i.e., f ∈ SK. Here K represents the set of possible values each coordinate in S can take. Thus, we can think of the image itself displayed on a fixed display as some function transforming the matrix of pixel space, possible values (colours) K, into a visual representation v. So, an image can be represented by some input f and some output from a function v(f). (Frery et al., 3)

This is an important starting point because this implies that our image data, which is represented as a matrix, will be closed under addition, subtraction, and multiplication, and thus will allow for us to use matrix algebra. We will find that because images may be different sizes (pixel by pixel dimensions) that the ability to change the pixel space, S (defined above), can be understood as dropping unnecessary variables, rather than some violation of the integrity of the data itself.

In what follows, we will attempt various ways to decrease the dimensionality, by identifying various subsets of our data, grouping that data together (using decision trees), fisher’s linear discriminant analysis, but before we can use that we want to group the pixels in some helpful way and then based on that grouping further decrease the dimensionality using a Linear discriminant analysis, and we will also use principle component analysis, and compare these supervised techniques to an unsupervised technique known as random forest. We will compare how each of these performs.

Before we move further, we must understand that ultimately it is the ability for a human to differentiate the concept of outdoors and not outdoors that is the basis upon which a correct answer is determined. So, it is useful to briefly describe how we come to as humans visually identify what is are essentially very abstract concepts, ‘outdoors’ and ‘not outdoors’.

Foundation:

The pixel space, which we plane to turn into the feature set is compact (Frery et al, 2). A compact space, by the definition of general topology, is a bounded probability space (The Editors of Encyclopedia Britannica, 2017). So, while we may be able to represent the presence of an outdoor image or a ‘not-outdoor image’ with by a binomial distribution, we cannot assume the same about the pixel space. The pixel space is bounded, there is a specific range of values it can take and a maximum set of possibilities, and it is not distributed necessarily across any real number space, so a Gaussian distribution, which would otherwise assign positive probabilities to ‘impossible’ images would not be appropriate for creating our feature set. However, as stated, once our feature set is created, we may be able to use a binomial distribution, as performed in the reference code, to classify our images.

One interesting and exciting consideration is the fact that we may be able to use one important distinction between sunlight and room lighting. That is, room lighting is created for the sake of the human eye, while sunlight is not. Therefore, room lighting, would not need to be manufactured to include light that humans cannot see, thus may have a spectrum of light that may appear in images with a resolution at or beyond the human eye (such as images with resolutions beyond 4k), or may have certain colour differences that the human eye cannot differentiate but may be present. So, this could be an opportunity for a machine learning algorithm such as a neural network, to take advantage of.

Work Cited:

Frery, Alejandro C., and Talita Perciano. *Introduction to Image Processing Using R: Learning by Examples*. Springer, 2013.

The Editors of Encyclopaedia Britannica. (2017, February 2). Compactness. Retrieved March 20, 2020, from https://www.britannica.com/science/compactness