**Pixel value distribution in dermoscopic images of skin lesions for melanoma detection**

Malignant melanoma of the skin (melanoma) is the deadliest type of skin cancer. Early detection is crucial since expected patient survival drops rapidly as the tumor evolves. Early stage melanomas resemble common skin lesions (moles), and are therefore difficult to detect. A dermoscope consists of a magnifying lens, a glass plate and surrounding light. The light of the dermoscope penetrates the uppermost skin layer and reveals structures that are invisible to the naked eye, and is therefore used for early detection of melanoma.

The colour of the lesion is an important feature for melanoma detection, since melanomas can have other colours in addition to the normal brown colour of a skin lesion.

Computer aided diagnostic (CAD) systems for melanoma detection follow the procedure of (i) image pre-processing, (ii) segmenting the lesion from the background skin, (iii) feature value calculation, (iv) feature selection, and (v) classification. There are commonly several colour feature algorithms in one system. The algorithms either aim at detecting melanoma specific colours (such as blue-grey), counting the number of colours (melanomas typically have more colours than benign lesions), and describing the colours statistically (e.g. mean and standard deviation of the observed pixel values).

We propose a new type of colour feature to discriminate between melanoma and benign skin lesions which estimates the pixel value distribution of a lesion and then measure its dissimilarity to a known distribution (benign or malignant).

The pixel value distribution of a lesion is estimated by sampling 1000 pixels from the lesion area of the image, and then fitting a Gaussian mixture model (GMM) to the pixel values. The number of components in the mixture is estimated by the Bayesian Information Criterion (BIC). The benign distribution is estimated by sampling 1000 pixels from a each lesion in a collection of benign lesion images, and then fitting a GMM. The malignant distribution is estimated equivalently.

The dissimilarity between a lesion distribution and the benign (malignant) distribution is measured by a weighted version of the Kullback-Leibler information. The Kullback-Leibler information is a measure of divergence, i.e. how “far away” two distributions are from each other. The weighting is performed to emphasise the regions of the colour space of melanoma specific colours, through using pixel values from melanoma images in the numerical approximation of the Kullback-Leibler intergral. We then have two features; d(les,ben) and d(les,mal) which reflect the dissimilarity between an unclassified lesion and the benign and the malignant model, respectively.

We have measured the new features performance by (i) their sensitivity and specificity for skin lesion classfication, (ii) how often they are picked in a pool with 59 other lesions by a feature selector, and (iii) their contribution to the performance of a skin lesion classifier.