AUTOMATED DOMAIN-SPECIFIC FEATURE SELECTION FOR CLASSIFICATION-BASED SEGMENTATION OF TOMOGRAPHIC MEDICAL IMAGE DATA

Gerald Zwettler[a,b,c], Werner Backfrieder[b]

[a]Bio- and Medical Informatics, Research and Development Department, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, AUSTRIA
[b]School of Informatics, Communication and Media, University of Applied Sciences Upper Austria, Softwarepark 11, 4232 Hagenberg, AUSTRIA
[c]Research Group Scientific Computing, Faculty of Computer Science, University of Vienna, Währingerstrasse 29/6.21, 1090 Vienna, AUSTRIA

[a]gerald.zwettler@fh-hagenberg.at, [b]werner.backfrieder@fh-hagenberg.at

ABSTRACT

Classification-based segmentation is an approach to establish generic analysis of medical image data. Significant feature sets covering different characteristics of regions to segment allow for robust discrimination of topologically defined classes. In this work a method for automated domain-specific feature selection to achieve a higher level of predictability is presented, incorporating multivariate feature analysis. For calculation of the probability density function, different approaches, like histogram analysis, enumeration of the entire feature space or umbrella Monte Carlo Integration are investigated. Furthermore, meta features calculated on entire classification results rather than on particular regions are introduced. Predictability of both, single local and meta features, is evaluated for different medical datasets as well for simulated intensity volumes, allowing testing and evaluating specific classification problems. The automated feature selection proofs to be accurate for classification-based segmentation utilizing well-known machine learning approaches.

Keywords: classification-based segmentation, multivariate feature analysis, Monte Carlo Integration, automated feature selection

1. INTRODUCTION

Precise segmentation of target anatomical structures from tomographic image datasets is an essential prerequisite for quantitative analysis and computer-assisted diagnostics. If all voxels of an anatomical structure are labelled, measurements on the extent and volume become feasible, for instance facilitating the monitoring of the disease progression. Furthermore, from available segmentations 3D surface models can be derived that can be utilized for surgery planning (Zwettler, Backfrieder, Swoboda, and Pfeifer 2009) or surgical training (Fürst and Schrempf 2012). Thereby the user interaction and subsequent analysis can be performed on the computer model or utilizing a virtual reality environment, enriched by haptic patient models that are derived from the anatomical segmentations and produced via emerging 3D printing devices. Precise segmentations are not only required for computer-based analysis, but also for registering multi-model image data of the same patient to combine high resolution morphological imaging (*CT*, *MRI*) and image data from the functional imaging domain (*PET*, *SPECT*). Only with available segmentation masks, the measured metabolic activity can be limited to organ borders to be quantitatively evaluated with respect to anatomical classifications (Beyer, Schwenzer, Bisdas, Claussen, and Pichler 2010).

In the last decades there has been intensive research work in the field of medical image processing to achieve preferably fully-automated segmentation approaches in specific diagnostic domains. Utilizing deformable models (McInerney and Terzopoulos 1996) and incorporating a priori knowledge on the target anatomical structure, morphologies with low variability in shape can be robustly segmented. Nevertheless, generic application of deformable models for arbitrary segmentation domains is not feasible as proper adjustment of the parameters and the a priori model is required. In contrast, Statistical Shape Models (Cootes, Taylor, Cooper, and Graham 1992) can be trained rather autonomously, if a large set of reference segmentations covering all relevant possible anatomical variations is available. Active appearance models (Cootes, Edwards, and Taylor 1998) introduce additional statistical properties of the targets structure expected intensity profile besides geometric features and Level sets (Osher and Sethian 1988) can handle changes in topology and anatomical variability but complex parameterization needs adjustment to the particular segmentation task. Furthermore, all of these sophisticated models are limited to segmentation of particular anatomical shapes, as border areas and overlapping segments cannot be handled, when segmenting multiple classes from input volume, covering all of the voxels at most.

Segmentation of the entire dataset from arbitrary imaging domains, i.e. assigning a class label to all available voxels, is up to now not feasible by utilizing

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

26

fully-automated model-based approaches. However, such entire segmentations can be achieved in a semi-automated way by utilizing conventional segmentation approaches like region growing (Gonzalez and Wintz 1987) or live wire contour detection (Schenk and Prause 2001) with appropriate filtering and morphological post-processing in a rapid prototyping image processing pipeline. A standardized process model for segmentation of arbitrary anatomical structures from variable tomographic image data has been presented in (Zwettler and Backfrieder 2013). While this approach is by far too user-intensive for practical application, the reference segmentations achievable in a semi-automated way are perfectly suited for training a priori models of specific segmentation domains.

A fully automated segmentation of all anatomical structures present from the particular input image modalities is achieved utilizing classification-based approaches, where at first fragmenting the input volume into regions of similar intensities, demarcated by gradients along the borders. For this pre-processing, e.g. watershed transform (Vincent and Soille 1991) applied to gradient magnitude can be utilized. To introduce additional robustness for this pre-segmentation, confidence-connected intensities, neighbourhood characteristics for iterative region merging and morphological post-processing are additionally incorporated (Zwettler and Backfrieder 2012). Given that pre-segmentation of the input image leads to fragmentation into mosaic-like regions with voxels predominantly belonging to one particular class, the actual segmentation can be achieved via multivariate feature classification. Therefore different characteristics of the particular anatomical structures, like intensity statistics, shape, location or co-occurrence metrics (Felipe, Traina and Traina 2003), are applicable. Most of these features are normally distributed, thus allowing classification with Gaussian mixture models, probability density function, Bayes networks or k-Means clustering. Furthermore, powerful classifiers from machine learning domain like neural networks (Vapnik 2000) or support vectors (Boser, Guyon and Vapnik 1992) are applicable. Besides, also heuristic approaches like genetic programming (Koza 1992) or genetic algorithms (Goldberg 1989) can be utilized for feature-based image classification.

In this work, $n=30$ local features are evaluated with respect to their predictability, i.e. achievable classification precision for specific objectives, in different medical imaging domains. Thereby, *statistical*, *geometric*, *texture* and *transformation* features are utilized. For many of the proposed classifiers, the number of feature dimensions has to be limited due to numeric computability, over-training and noise sensitivity. Thus, in this work an approach is developed to automatically select the most appropriate feature set for a specific medical segmentation domain, according to single feature predictability considering feature correlation. Thereby, also the classification targets like high voxel match, high region classification confidence

or individual class weights can be incorporated. To introduce additional robustness and to allow classification in a broader range of application, additional $m=13$ meta features are introduced that are calculated on the class region statistics. These meta features are perfectly suited to be utilized for heuristic classification. For automated features selection, different cumulated predictability metrics, like multivariate PDF or histogram analysis are evaluated.

For testing and validation of the proposed features and proposed domain-specific feature selection, different medical datasets and simulated intensity volumes (Zwettler and Backfrieder 2014) are utilized.

2. MATERIAL
For testing and validation, MRI data from *BrainWeb* database (Cocosco, Kollokian, Kwan and Evans, 1997), anonymized patient studies from various imaging modalities and simulated intensity datasets are utilized as detailed in the following section.

2.1. Data from *BrainWeb*
In total $n=20$ MRI datasets, reflecting spoiled FLASH sequences at flip angle $\alpha=30$ and an echo time (TE) of *9.2ms* and a repetition time (TR) of *22ms* with associated reference segmentations are utilized. The test sequence datasets denoted as *BRAINWEB_REF* in the following, have *0.5mm* voxel spacing at volume dimensions of *256×256×181* and *unsigned 8bit* scalar range. Reference segmentations fragment the dataset according to a topology of $k=12$ classes reflecting anatomical structures to discriminate, namely *background* (0), *cerebrospinal fluid* (1), *grey matter* (2), *white matter* (3), *fat* (4), *muscles* (5), *skin* (6), *skull* (7), *vessels* (8), *around fat* (9), *dura mater* (10) and *bone marrow* (11).

Inter dataset class ratio variability of *BRAINWEB_REF* sequence is charted in Fig. 1, while axial slices #50 and #100 of *BrainWeb* dataset *ds1* are presented in Fig. 2.
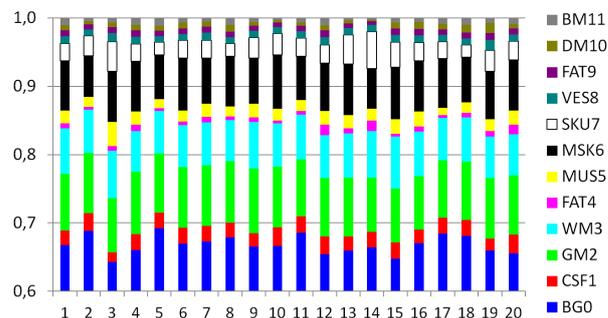
Figure 1: Inter dataset variability of the $n=20$ datasets from *BREANWEB_REF*. Voxel ratio of the particular classes C_i significantly varies.

For *BrainWeb* datasets moreover manual reference segmentations are prepared as sequence *BRAINWEB_MAN*, discriminating between $k=6$ classes, namely *white matter* (0), *grey matter* (1), *ventricle* (2), *background* (3), *tissue* (4) and *remaining* voxels (5).

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

27

Surface renderings for reference segmentations of *ds1* are presented in Fig. 3.

(a)   (b)

(c)   (d)
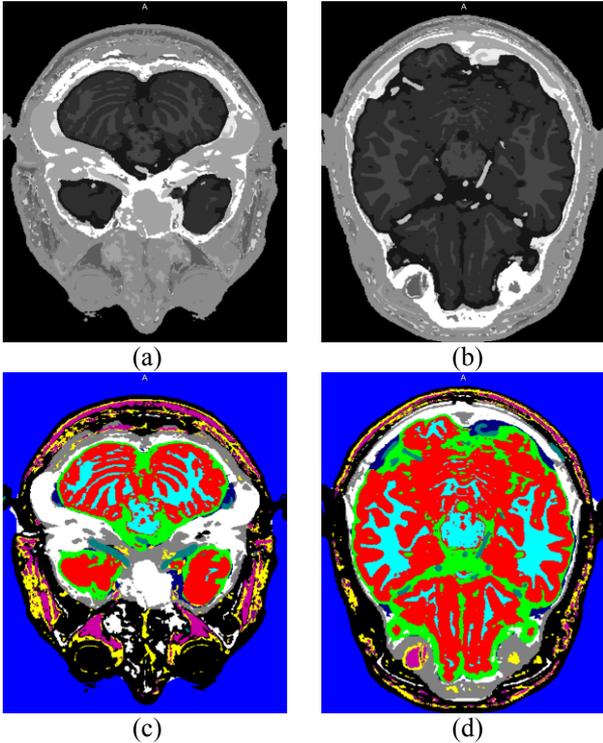
Figure 2: First *BREANWEB_REF* dataset with input intensities and reference class labels.

(a)   (b)   (c)
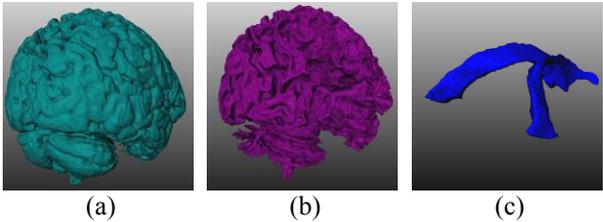
Figure 3: Surface Renderings on *BRAINWEB_MAN ds1* for grey matter (a), white matter (b) and ventricle (c).

2.2. Anonymized Patient Studies

Preparation of reference segmentation for true patient datasets necessitates the use of standardized image processing chains (Zwettler and Backfrieder 2013).

The sequence *HEART* thereby covers *n=15* datasets acquired with Siemens Somatom Sensation Cardiac 64-row MSCT, showing *512×512* slice dimensionality and *μ=288.5(92-461)* slices with an average spacing of *.359* in x/y direction and *.56* in z-direction. The defined topology discriminates between *k=9* classes, namely *lung* (0), *aorta* (1), *left ventricle* (2), *stents* (3), *right ventricle* (4), *liver* (5), *bones* (6), *tissue* (7) and *remaining* voxels (8).

Furthermore, abdominal CT datasets are utilized to build up sequence *ABDOMEN*, fragmenting input voxels into *k=13* classes, namely *background* (0), *intestinal tract* (1), *lungs* (2), *muscles* (3), *aorta* (4), *kidneys* (5), *stomach* (6), *vessels* (7), *liver* (8), *heart* (9), *bones* (10), *tissue* (11), and *remaining* voxels (12).

An overview of sequence *HEART* is given in Fig. 4 for *ds1* with axial slice #20 in (a) and labelled reference segmentations in (b) with surface renderings of the heart ventricles and aorta in (c). For *ds1* of *ABDOMEN* sequence, axial slice #50 is shown in (d) and reference segmentations in (e) with surface renderings of lungs, aorta and liver in (f).

(a)   (b)   (c)

(d)   (e)   (f)
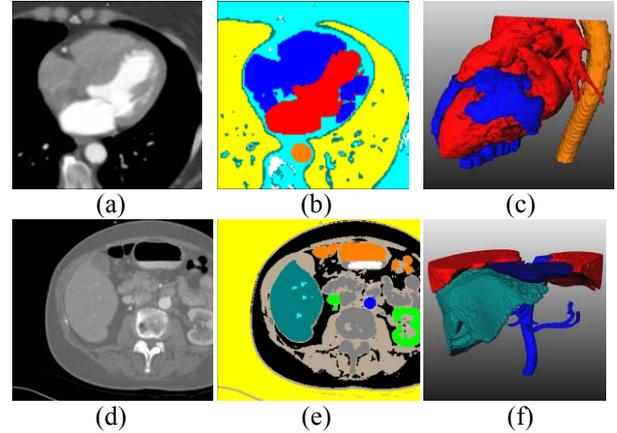
Figure 4: Overview of *HEART* and *ABDOMEN* data.

2.3. Simulated Intensity Volumes

A newly developed intensity volume simulator (Zwettler and Backfrieder 2014) allows for generation of testing sequences with different shape morphology of the particular classes C_i. Besides, intra region intensity characteristics, inter dataset variability, surface characteristics, region size, voxel ratio and several more attributes can be parameterized for simulation.

Simulated test sequences with *n=15* datasets are generated at mask dimensionality of *128×128×128* and allow for simulation of *tubular*, *BLOB*-like, *plane*-like and *scatter* morphologies. Sequence *SIM_1* thereby features discrimination of *k=5* classes, see Fig. 5.
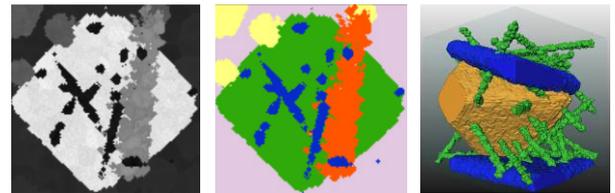
Figure 5: axial slice #50 of *SIM_1 ds1* (a) and associated reference segmentations (b) with surface rendering, showing tubular, plane and BLOB shape.

The testing sequences *SIM_2-SIM_8* only comprise two classes C_2 and C_3 with BLOB morphology showing identical characteristics that are differentiated from background class C_1. For the particular testing sequences, only specific characteristics are varied to allow for classification of classes C_2 and C_3 with respect to specific local and meta features, see Table 1.

3. METHODOLOGY

Whenever analysing statistically varied values of feature F_j calculated from a set of regions $D_i = \{R_1, R_2, \ldots, R_n\}$ assigned to a class C_i, the random feature values

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

28

generally can be modelled as normal distribution with $F_{ji} \sim N(\mu_{ji}, \sigma^2_{ji})$. Based on the single feature distributions, predictability at different positions in the multi-dimensional feature space and similarity evaluation with particular classes becomes feasible.

Table 1: Simulation Characteristics for *SIM_2-SIM_8*

ID	difference of C_2 and C_3
2	no difference ➔ utilized as ground truth
3	varied position within dataset
4	small differences in intensity profile
5	difference in connectedness, 1 vs. 60 islands
6	difference in mean region size
7	difference in class voxel ratio
8	intra region scalar correlation varied

3.1. Local and Meta Features

The *n=30* local features $\mathbf{F}_j$ considered in this work are calculated on particular regions R_i leading to static results, while the *m=13* introduced meta features $\hat{\mathbf{F}}$ are calculated on the classified region set D_i of a particular class C_i, allowing to incorporate single region classification results.

The total *43* considered features are grouped into four categories, namely *texture* features (Table 3), *statistical* features (Table 2), *geometric* features (Table 4) and *transformation* features (Table 5), covering different aspects on how to robustly discriminate the particular classes C_i.

With the *texture* features, as presented in Table 3, key intensity profile characteristics like mean region intensity, energy or entropy are incorporated.

The *statistical* features in Table 2 are mainly meta features, which evaluate statistics on region size calculated over all regions assigned to a particular class. Besides, class size and voxel ratio are calculated.

The *18 geometric* features enlisted in Table 4 refer to positional aspects of the particular regions, like mean position in x-, y-, and z-direction as well as mean city block distance from the image centre. Besides, surface-to-volume ratio and sphericity are calculated to address compactness.

Finally, for calculation of the *5 transformation* features from Table 5, co-occurrence matrix is derived from input image to calculate variance, entropy, energy and homogeneity. With these features, intensity profile characteristics can be well described.

Table 2: *statistical* features (9)

ID	definition	description
F_{12}	$F_{12}(\mathcal{R}) = |\mathcal{R}|$	region size
$\hat{F}_{13}$	$\hat{F}_{13}(\mathfrak{R}_{X_o}) = \min F_{12}(\mathfrak{R}_{X_o})$	min region size for class $\mathcal{X}_o$
$\hat{F}_{14}$	$\hat{F}_{14}(\mathfrak{R}_{X_o}) = \max F_{12}(\mathfrak{R}_{X_o})$	max region size for class $\mathcal{X}_o$
$\hat{F}_{15}$	$\hat{F}_{15}(\mathfrak{R}_{X_o}) = mean F_{12}(\mathfrak{R}_{X_o})$	mean region size for class $\mathcal{X}_o$
$\hat{F}_{16}$	$\hat{F}_{16}(\mathfrak{R}_{X_o}) = quantile(0.5, F_{12}(\mathfrak{R}_{X_o}))$	median region size for class $\mathcal{X}_o$
$\hat{F}_{17}$	$\hat{F}_{17}(\mathfrak{R}_{X_o}) = quantile(0.25, F_{12}(\mathfrak{R}_{X_o}))$	quantile 25 region size for $\mathcal{X}_o$
$\hat{F}_{18}$	$\hat{F}_{18}(\mathfrak{R}_{X_o}) = quantile(0.75, F_{12}(\mathfrak{R}_{X_o}))$	quantile 75 region size for $\mathcal{X}_o$
$\hat{F}_{19}$	$\hat{F}_{19}(\mathfrak{R}_{X_o}) = std F_{12}(\mathfrak{R}_{X_o})$	σ in region size for $\mathcal{X}_o$
$\hat{F}_{20}$	$\hat{F}_{20}(\mathfrak{R}_{X_o}) = \frac{1}{|I|} \cdot \sum_{i=1}^{|\mathfrak{R}_{X_o}|} |\mathcal{R}_i|$	component size ratio of $\mathcal{X}_o$

Table 3: *texture* features (11)

ID	definition	description
F_1	$F_1(\mathcal{R}) = mean\, I(\mathcal{R})$	mean region intensity value
F_2	$F_2(\mathcal{R}) = \min I(\mathcal{R})$	minimum intensity in $\mathcal{R}$
F_3	$F_3(\mathcal{R}) = \max I(\mathcal{R})$	maximum intensity in $\mathcal{R}$
F_4	$F_4(\mathcal{R}) = std\, I(\mathcal{R})$	standard deviation of intensity values in region $\mathcal{R}$
F_5	$F_5(\mathcal{R}) = quantile(0.5, I(\mathcal{R}))$	median intensity value in region $\mathcal{R}$
F_6	$F_6(\mathcal{R}) = quantile(0.25, I(\mathcal{R}))$	quantile 25 intensity value in region $\mathcal{R}$
F_7	$F_7(\mathcal{R}) = quantile(0.75, I(\mathcal{R}))$	quantile 75 intensity value in region $\mathcal{R}$
F_8	$F_8(\mathcal{R}) = \sum_{i=s_{min}}^{s_{max}} p_i \cdot \log p_i$	entropy of intensities
F_9	$F_9(\mathcal{R}) = \sum_{i=s_{min}}^{s_{max}} p_i^2$	energy of intensities
F_{10}	$F_{10}(\mathcal{R}) = \max_{i=s_{min}}^{s_{max}} p_i$	maximum probability of intensities
F_{11}	$F_{11}(\mathcal{R}) = \frac{1}{s_{max}-s_{min}} \cdot \sum_{i=s_{min}}^{s_{max}} p_i$	mean probability of intensities

Table 4: *geometric* features (18)

ID	definition	description
F_{21}	$F_{21}(\mathcal{R}) = \frac{surf(\mathcal{R})}{|\mathcal{R}|}$	surface to volume ratio
F_{22}	$F_{22}(\mathcal{R}) = \psi(\mathcal{R})$	sphericity
F_{23}-F_{25}	$F_{23}(\mathcal{R}) = C_{oM}(\mathcal{R})_x$	region center of mass for x-coordinate (F_{23}), y-coordinate (F_{24}) and z-coordinate (F_{25})
F_{26}	$F_{26}(\mathcal{R}) = distCB(C_{oM}(\mathcal{R}), C_{oM}(\mathfrak{R}_o))$	city block distance of region center-of-mass to expected position
F_{27}-F_{29}	$F_{27}(\mathcal{R}) = \frac{1}{|\mathcal{R}|}\sum_{i=1}^{|\mathcal{R}|} |c_{i_x} - C_{oM}(\mathcal{R})_x|$	mean city-block distance from center-of-mass in x direction (F_{27}), y direction (F_{28}) and z direction (F_{29})
F_{30}	$F_{30}(\mathcal{R}) = \frac{1}{|\mathcal{R}|}\sum_{i=1}^{|\mathcal{R}|} distCB(c_i, C_{oM}(\mathcal{R}))$	mean city block distance to center-of-mass
F_{31}-F_{33}	$F_{31}(\mathcal{R}) = \frac{1}{|\mathcal{R}|}\sum_{i=1}^{|\mathcal{R}|} |c_{i_x} - C_{oM}(I)_x|$	mean city-block distance from image I center-of-mass in x direction (F_{31}), y direction (F_{32}) and z direction (F_{33}).
F_{34}	$F_{34}(\mathcal{R}) = \frac{1}{|\mathcal{R}|}\sum_{i=1}^{|\mathcal{R}|} distCB(c_i, C_{oM}(I))$	mean city block distance to image I center-of-mass
$\hat{F}_{35}$	$\hat{F}_{35}(\mathfrak{R}_{X_o}) = numOfIslands(\mathfrak{R}_{X_o}, \mathcal{N}_{18})$	expected number of autonomous islands
$\hat{F}_{36}$-$\hat{F}_{39}$	$\hat{F}_{36}(\mathfrak{R}_{X_o}) = C_{oM}(\mathfrak{R}_{X_o})_x$	center-of-mass calculated for all regions of class $\mathcal{X}_o$ in x direction ($\hat{F}_{36}$), y direction ($\hat{F}_{37}$) and z direction ($\hat{F}_{38}$)

Table 5: *transformation* features (5)

ID	definition	description
F_{39}	$F_{39}(\mathcal{C}_{\mathcal{R}}) = \sum_{i=s_{min}}^{s_{max}}\sum_{j=s_{min}}^{s_{max}} (i-j)^2 \cdot p(i,j)$	variance of region
F_{40}	$F_{40}(\mathcal{C}_{\mathcal{R}}) = \sum_{i=s_{min}}^{s_{max}}\sum_{j=s_{min}}^{s_{max}} p(i,j) \cdot \log p(i,j)$	entropy of region
F_{41}	$F_{41}(\mathcal{C}_{\mathcal{R}}) = \sum_{i=s_{min}}^{s_{max}}\sum_{j=s_{min}}^{s_{max}} p(i,j)^2$	energy of region
F_{42}	$F_{42}(\mathcal{C}_{\mathcal{R}}) = \sum_{i=s_{min}}^{s_{max}}\sum_{j=s_{min}}^{s_{max}} \frac{p(i,j)}{1+|i-j|}$	homogeneity of region
F_{43}	$F_{43}(\mathcal{C}_{\mathcal{R}}) = \max p(i,j)$	maximum value as max co-occurrence

3.2. Single Feature Predictability

Features $\mathbf{F}_j$ showing a high level of predictability are considered to be best suited to discriminate the considered classes C_i. As the normal distributions of the particular classes are expected to at least partially overlap, classification generally cannot be performed

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

29

without bearing some level of uncertainty, see Fig. 6 for illustration of one-dimensional feature set with distributions for classes $C_1 - C_3$ partially overlapping.
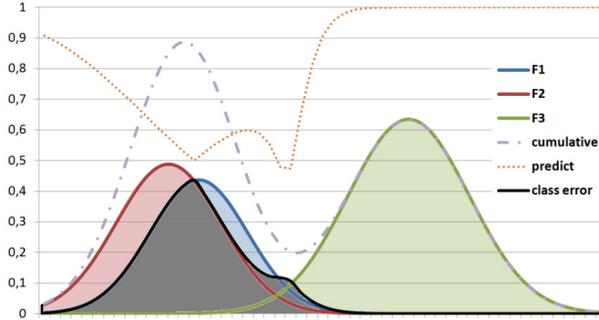
Figure 6: Bayes error in feature classification. The overlapping area marked in black of the class feature distributions F_1, F_2 and F_3 derived from classes C_1, C_2 and C_3 for feature F_j is classified at low confidence with $pred_j=.771$ as F_1 and F_2 are largely overlapping.

The local probability density function (*pdf*) of the particular class feature value distributions is thereby defined for a particular class X_o as

$$pdf(x, \mathcal{X}_o, \mathcal{F}_{l_o}) = \frac{1}{\sqrt{2\pi\sigma_{l_o}^2}} \exp\left(-\frac{(x-\mu_{l_o})^2}{2\sigma_{l_o}^2}\right) \quad (1)$$

where the pdf is weighted according to region occurrence ratio, voxel class ratio or utilizing specific weight for the particular classes in the sense of Bayesian inference as delineated in Equ. 2. This way, the particular key classification objective is defined and differences in region or voxel occurrence probability of the particular classes can be modelled.

$$P(\mathcal{X}_o|x) = pdf(f, \mathcal{X}_o, \mathcal{F}_{l_o}) \cdot P(\mathcal{X}_o) \quad (2)$$

Predictability (*pred*) is evaluated at each discrete position of the feature space with

$$pred(y, \mathcal{F}_{l_o}) = \frac{\max\limits_{\mathcal{X}_i \in \mathfrak{x}} pdf(y, \mathcal{X}_i, \mathcal{F}_{l_o}) \cdot P(\mathcal{X}_i)}{\sum\limits_{\mathcal{X}_i \in \mathfrak{x}} pdf(y, \mathcal{X}_i, \mathcal{F}_{l_o}) \cdot P(\mathcal{X}_i)} \quad (3)$$

so that overall predictability of a particular feature with respect to all considered classes X_i can be formulated as

$$pred(\mathcal{F}_l) = \frac{\int_{y=-\infty}^{\infty} \max\limits_{\mathcal{X}_i \in \mathfrak{x}} pdf(y, \mathcal{X}_i) \cdot P(\mathcal{X}_i)}{\int_{y=-\infty}^{\infty} \sum\limits_{\mathcal{X}_i \in \mathfrak{x}} pdf(y, \mathcal{X}_i) \cdot P(\mathcal{X}_i)}. \quad (4)$$

3.3. Multivariate Class Similarity

For common classification tasks the sole utilization of single features generally is considered to be insufficient. Instead, several features are utilized to derive the best matching class label of a region R_i from the particular feature vector. With every additionally incorporated feature, the dimensionality of the feature space is increased by one, thereby increasing the overall achievable predictability, see Fig. 7 for two-dimensional

feature space with distributions for classes $C_1 - C_3$ less overlapping compared to each single feature dimension.
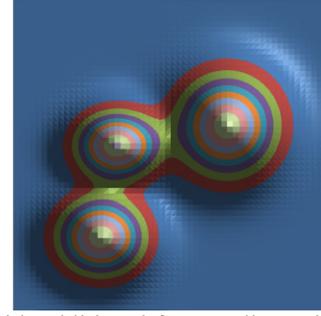
Figure 7: With additional feature dimension F_k showing $pred_k=.743$ to be combined with F_j from Fig. 6, overall predictability is increased to $pred_{jk}=.9707$.

For calculation of cumulated predictability from single feature dimensions, correlation has to be considered. With higher feature correlation τ_{ij}, of the involved features F_j and F_k, the increase in overall predictability due to the extended feature space drops.
Calculation of multi-dimensional probability density function pdf_{gmm} is achieved with

$$pdf_{gmm}(X, \mathcal{X}_o, \mathfrak{F}_o) = \frac{1}{(2\pi)^{d/2} \cdot \sqrt{|\Sigma|}} \cdot$$
$$exp\left(-\frac{1}{2}(X - \mu_{\mathfrak{F}_o})' \Sigma^{-1}(X - \mu_{\mathfrak{F}_o})\right) \quad (5)$$

for input vector $X = \{x_1, \ldots, x_d\} \in \mathbb{R}^d$ with covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ calculated for pairwise covariances as

$$COV(X_k, X_l) = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_{k_i} - \bar{X}_k) \cdot (x_{l_i} - \bar{X}_l) \quad (6)$$

3.4. Automated Domain-Specific Feature Selection

With the formulations defined in section 3.3, predictability of a certain feature set can be calculated for particular segmentation domains by statistically evaluating a set of reference regions $D_i = \{R_1, R_2, \ldots, R_n\}$ preserved for each class C_i to be discriminate. Nevertheless, with increasing dimensionality of the feature space, the proper calculation strategy for integration over the feature space from local evaluation of the pdf_{gmm} has to be analysed in depth. Therefore, the following calculation strategies are applicable, namely:

- Partial integration by evaluating pdf_{gmm} at discrete and equally distributed positions per feature dimension is only applicable for a very small feature vector size, as the total number of required evaluation positions exponentially grows with respect to increased feature space dimensionality. A reduction of sampling positions per dimension would allow for handling higher-dimensional feature spaces but entail a drop in result quality as the sampling frequency becomes deficient.
- *Monte Carlo integration* (*MCI*) (Metropolis 1987) allows for sampling of the feature space at a predefined number of random positions. This way,

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

30

also higher dimensions can be handled. Nevertheless, insufficient sampling frequency as mentioned for the partial integration approach, remains unaddressed. With decreasing sampling density, the risk for leaving significant parts of the feature space unconsidered during MCI increases. Imprecision introduced by MCI is especially high when single class distribution show small variability and thus might not be hit by any of the sampling positions at all or if the feature space is generally sparsely populated.

- The *umbrella* variation of *Monte Carlo* integration addresses the problem of feature space sparseness by sampling only at positions derived from the particular class feature value deviations, incorporating Bayes inference with respect to region or voxel class probabilities.

- Similar results can be achieved, when evaluating the pdf_{gmm} exclusively at positions of the multi-dimensional feature space that are derived from the incorporated set of reference regions. Nevertheless, applicability of this approach highly depends on a sufficient number of reference segmentation datasets. Furthermore, inaccuracies are introduced from feature distributions that do not perfectly approach Gaussian shape, as the pdf_{gmm} sampling is performed according to true region feature data.

- Finally, a *histogram* approach can be utilized, evaluating a multi-dimensional histogram for all classes C_i. This way, the true feature value distribution of the particular classes is approximated best. Thus, higher accuracy for the evaluated predictability values is to be expected for those features, e.g. for geometric properties, that do not perfectly approach Gaussian shape. To overcome the problem of features space sparseness, filtering is applied with respect to the particular bin count per dimension, which is downscaled with higher dimensionality, see Table 6.

Table 6: Histogram calculation parameters with respect to dimensionality $\mathcal{D}$

$\mathcal{D}$	smoothing filter kernel size k	histogram bins bin size s	total $s^{\mathcal{D}}$
1	10	200	200
2	8	100	10,000
3	5	50	125,000
4	3	20	160,000
5	2	10	100,000
6	1	8	262,144
7	0	6	279,936
8	0	5	390,625
9	0	4	262,144
10	0	3	59,049

Despite the different enlisted calculation strategies applicative to multi-feature predictability estimation, the algorithm for automated domain-specific feature selection remains the same. For automated selection of a feature set $\mathcal{F}_{local}$ and a feature set $\mathcal{F}_{meta}$ for a particular classification domain, defined by class-labelled regions R_i derived from the preserved set of reference segmentations, at first the feature with highest *pdf* is selected. In the following, the feature vector is iteratively extended until the target number of dimensions is approached, by selecting the feature $\mathbf{F}_j$ that maximizes the pdf_{gmm} achievable by the particular feature vector length, at each iteration.

Although in theory the overall achievable predictability pdf_{gmm} is expected to monotonically increase with every additional feature dimension $\mathbf{F}_j$, depending on the calculation strategy also decreasing overall predictability values might be noticed. This effect results from the smaller histogram bin size at higher dimensionality, which subsequently leads to a collapse of neighbouring bins and thus a pretended reduction in predictability. The same holds for *MCI* approaches, where the sampling frequency drops at increased dimensionality. Thus, any possible drop in overall predictability pdf_{gmm} thereby only results from the numerical calculation strategy, not due to effectively reduced classification capability. Nevertheless, for the actual classification problem, the chosen iterative feature selection approach allows for best feature selection with an ever increased predictability from each single feature added to the feature space.

4. IMPLEMENTATION

The algorithms developed for this work in the field of feature analysis and feature selection are implemented in *C++*. For I/O operations and random number generation, *boost* library (Boost 2014) is utilized as external library. Although, the core boost concepts are meanwhile added to *C++11* standard, the utilization of boost programming paradigms started with the development of the *MIPP* framework (Swoboda et al. 2008) back in 2007 was perpetuated for reasons of consistency. Besides, no further external code was utilized.

4.1. Image Volume Data Layout

The data layout for image volumes is a crucial aspect with respect to genericity and performance. For this work, 2D images and 3D volumes are represented as real 2D and 3D matrices of scalar values respectively.

Commonly, image processing frameworks like ITK (Kitware 2014) feature a one-dimensional data layout for images of arbitrary dimensionality, ensuring a maximum level of genericity. Nevertheless, with required index conversions from 1D to n-D at each pixel access, this design would not feature the excessive use of mask operations as required for gradient calculation, common filtering and morphological opening/closing operations.

Thus, 2D and 3D matrices of scalar values are preferably utilized, allowing direct access to the particular pixel/voxel positions via index tuples. The chosen approach does not show significant disadvantages with respect to memory management

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

31

(paging, memory allocation,...), and thus perfectly meets the procedural algorithmic requirements.

4.2. Histogram Data Layout

For histogram analysis in the context of multivariate feature analysis, also volume data of higher (at least 4D and 5D) has to be considered. Consequently, for this particular domain a higher level of genericity is required. Thus, only for histograms and the required associated operations, like filtering, a 1D layout for the voxel data with index conversions for access is applied.

5. RESULTS

In the following, first single feature predictability is analysed prior to addressing multidimensional feature spaces and automated domain-specific feature selection.

5.1. Results on Single Feature Predictability

At first, single feature predictability is evaluated for the $n=30$ local and $m=13$ meta features evaluated on $k=12$ classes of *BRAINWEB_REF* dataset series, with predictability results for the considered local features F_j charted in Fig. 8 and the meta features $\hat{F}$ presented in Fig. 9. The predictability is thereby calculated with partial integration (P_{stat}), from reference region feature vectors (P_{reg}), utilizing MCI (P_{mc}), based on reference region histogram (P_{hist}) and based on umbrella MCI (P_{mcu}), with *pdf* weighting chosen to maximize the number of correct classified regions in contrast to a maximized number of correctly classified voxel. Despite the chosen predictability calculation strategy, the particular local features show a similar trend.
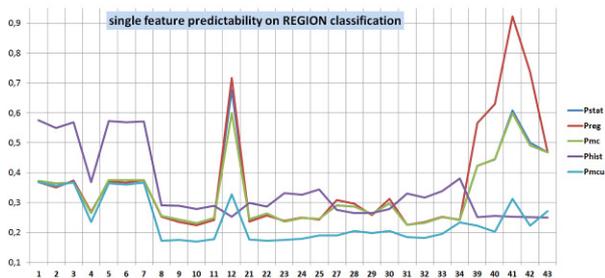
Figure 8: Predictability of local features F_k calculated for $n=20$ *BRAINWEB_REF* datasets.
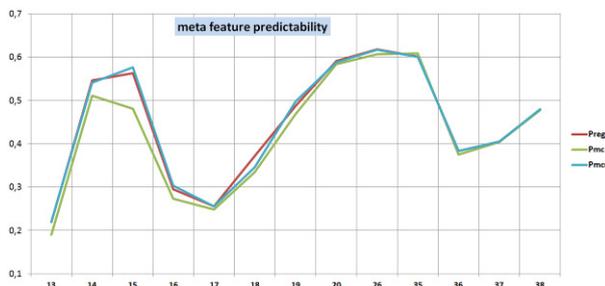
Figure 9: Predictability of meta features $\hat{F}$ calculated for $n=20$ *BRAINWEB_REF* datasets.

Besides single feature predictability, also feature correlation is relevant for later to follow multivariate feature analysis. Correlation matrix calculated for the first $k=15$ local features is presented in Fig. 10. As expected, the region intensity features (F_1-F_7) show a high level of correlation. The same holds for texture metrics directly calculated on image intensity profile (F_8-F_{11}). Nevertheless, with the different feature groups considered, enough feature independence is available for multi-dimensional feature analysis at high level of achievable predictability.

	1	2	3	4	5	6	7	8	9	10	11	12	21	22	23	24	25	27	28	29
1		0,98	0,98	0,33	1,00	1,00	1,00	0,14	0,16	0,06	0,20	0,03	0,16	0,18	0,02	0,08	0,03	0,03	0,06	0,12
2	0,98		0,95	0,20	0,98	0,99	0,97	0,24	0,25	0,15	0,28	0,03	0,23	0,26	0,02	0,08	0,01	0,07	0,09	0,17
3	0,98	0,95		0,46	0,98	0,97	0,99	0,05	0,08	0,01	0,12	0,01	0,10	0,11	0,02	0,08	0,06	0,02	0,01	0,07
4	0,33	0,20	0,46		0,31	0,25	0,38	0,22	0,21	0,28	0,16	0,03	0,02	0,08	0,01	0,02	0,15	0,01	0,02	0,01
5	1,00	0,98	0,98	0,31		0,99	0,99	0,13	0,15	0,06	0,19	0,03	0,15	0,17	0,02	0,08	0,03	0,03	0,05	0,12
6	1,00	0,99	0,97	0,25	0,99		0,99	0,15	0,17	0,08	0,20	0,03	0,16	0,18	0,02	0,08	0,02	0,03	0,05	0,12
7	1,00	0,97	0,99	0,38	0,99	0,99		0,11	0,13	0,03	0,17	0,03	0,15	0,16	0,02	0,08	0,04	0,03	0,05	0,11
8	0,14	0,24	0,05	0,22	0,13	0,15	0,11		0,94	0,81	0,95	0,01	0,72	0,85	0,01	0,01	0,01	0,23	0,23	0,35
9	0,16	0,25	0,08	0,21	0,15	0,17	0,13	0,94		0,90	0,98	0,02	0,64	0,75	0,00	0,01	0,01	0,16	0,17	0,27
10	0,06	0,15	0,01	0,28	0,06	0,08	0,03	0,81	0,90		0,81	0,02	0,16	0,18	0,02	0,08	0,02	0,03	0,13	0,14	0,22
11	0,20	0,28	0,12	0,16	0,19	0,20	0,17	0,95	0,98	0,81		0,03	0,70	0,79	0,00	0,01	0,03	0,21	0,23	0,32
12	0,03	0,03	0,01	0,03	0,03	0,03	0,03	0,01	0,02	0,02	0,03		0,15	0,00	0,00	0,00	0,01	0,67	0,67	0,48
21	0,16	0,23	0,10	0,02	0,15	0,16	0,15	0,72	0,64	0,50	0,70	0,15		0,49	0,00	0,04	0,03	0,34	0,38	0,39
22	0,18	0,26	0,11	0,08	0,17	0,18	0,16	0,85	0,75	0,59	0,79	0,00	0,49		0,01	0,03	0,08	0,32	0,34	0,49
23	0,02	0,02	0,02	0,01	0,02	0,02	0,02	0,01	0,00	0,00	0,00	0,00	0,00	0,01		0,03	0,01	0,00	0,00	0,01
24	0,08	0,08	0,08	0,02	0,08	0,08	0,08	0,01	0,01	0,01	0,01	0,00	0,04	0,03	0,03		0,12	0,04	0,04	0,03
25	0,03	0,01	0,06	0,15	0,03	0,02	0,04	0,01	0,01	0,01	0,03	0,01	0,03	0,08	0,01	0,12		0,04	0,04	0,00
27	0,03	0,07	0,02	0,01	0,03	0,03	0,03	0,23	0,16	0,13	0,21	0,67	0,34	0,32	0,00	0,04	0,04		0,78	0,69
28	0,06	0,09	0,01	0,02	0,05	0,05	0,05	0,23	0,17	0,14	0,23	0,67	0,38	0,34	0,00	0,04	0,04	0,78		0,75
29	0,12	0,17	0,07	0,01	0,12	0,12	0,11	0,35	0,27	0,22	0,32	0,48	0,39	0,49	0,01	0,03	0,00	0,69	0,75	

Figure 10: Correlation matrix calculated for the first $k=15$ local features F_k evaluated on *BRAINWEB_REF*.

The simulated dataset series *SIM_2-SIM_8* are perfectly suited to evaluate the local and meta features showing highest predictability for the specific classification scenarios. Results on the $k=30$ local features are presented in Fig. 11 and for the meta features in Fig. 12., with single feature predictability reported relative to results on *SIM_2* as ground truth. Thus, a peak indicates, that for the particular scenario, the particular feature F_j is at least superior utilizable for classification.
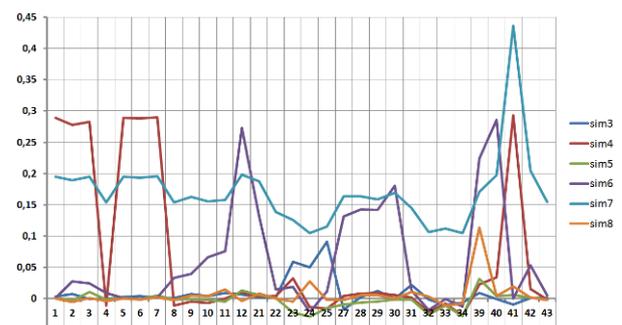
Figure 11: Single local feature predictability calculated for *SIM_3-SIM_8* relative to *SIM_2* as neutral classification scenario.

Analysis on the simulated datasets SIM_3-SIM_8 leads to the following findings:

- For *SIM_3* with only class position being varied, geometric features (F_{23}-F_{25}) are best.
- For *SIM_4* with varied mean intensity level, intensity features (F_1-F_7) contribute most for overall predictability.

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

32

- Meta feature $\hat{F}_{35}$ best addresses the difference in connectedness according to number of islands for *SIM_5*.
- For *SIM_6* the mean region size (F_{12}) and for *SIM_7* voxel-ratio meta feature $\hat{F}_{20}$ are best suited, whereas the intra region intensity characteristics simulated with *SIM_8* necessitate utilization of co-occurrence features (F_{39}-F_{43}).
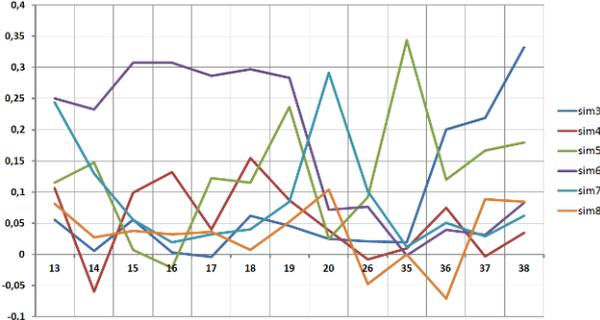
Figure 12: Single meta feature predictability calculated for *SIM_3-SIM_8* relative to *SIM_2* as neutral classification scenario.

Generally, predictability of the particular features shows a general tendency despite the particular segmentation domain and an overall high level, which correlates indirectly with the number of defined classes *k* in the particular classification domains, see Fig. 13.
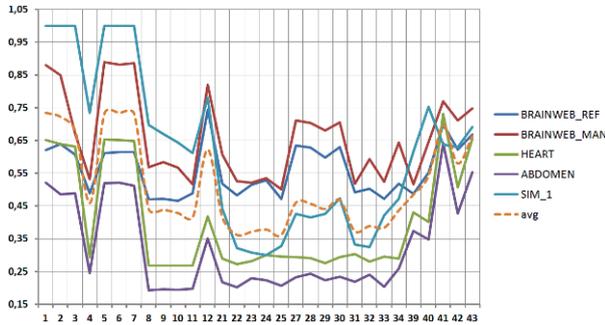
Figure 13: Single meta feature predictability calculated for *BRAINWEB_REF*, *BRAINWEB_MAN*, *HEART*, *ABDOMEN* and *SIM_1*.

5.2. Results on Multivariate Class Similarity and Domain-Specific Feature Selection

Although the particular local features show a common trend for the different segmentation and classification domains as previously discussed in Fig. 13, automated feature selection leads to different results for the utilized testing sequences as shown in Table 7. While for all of the considered datasets, feature F_6 is chosen at first rank, the features selected at rank F_{II} to F_V highly depend on the particular classification domain. Feature F_6 thereby best handles outliers compared to the other similar intensity features. Automated feature selection based on P_{hist} as shown in Table 7 generally leads to an increased overall predictability with every additional feature, while due to calculation reasons, the results at higher dimensionality also quantitatively drop due to

collapsing bins, see Fig. 14. In contrast, utilizing P_{mcu}, the cumulated results show monotonic increase also at higher dimensions, see Fig. 15.

Table 7: Domain-specific feature selection based on histogram predictability P_{hist}.

dataset		$\mathcal{F}_I$	$\mathcal{F}_{II}$	$\mathcal{F}_{III}$	$\mathcal{F}_{IV}$	$\mathcal{F}_V$
B_REF	$\mathcal{F}_i$	6	33	34	21	29
	P_{hist}	.99171	.90909	.98315	.89093	.93510
	$P_{total_{hist}}$	.99171	.99378	.99287	.98570	.97232
B_MAN	$\mathcal{F}_i$	6	8	34	40	29
	P_{hist}	.97430	.92915	.92717	.92342	.92069
	$P_{total_{hist}}$	.97430	.98281	.98335	.97645	.95609
HEART	$\mathcal{F}_i$	6	25	24	23	21
	P_{hist}	.79306	.48754	.52078	.50716	.44183
	$P_{total_{hist}}$	.79306	.81167	.82332	.80702	.76524
ABDOM	$\mathcal{F}_i$	6	29	23	24	25
	P_{hist}	.71942	.48264	.48561	.46589	.45236
	$P_{total_{hist}}$	.71942	.73392	.75085	.73710	.70514
SIM_1	$\mathcal{F}_i$	6	40	1	8	2
	P_{hist}	.99956	.82879	.99938	.80552	.99945
	$P_{total_{hist}}$	.99956	.99897	.99763	.98726	.97692
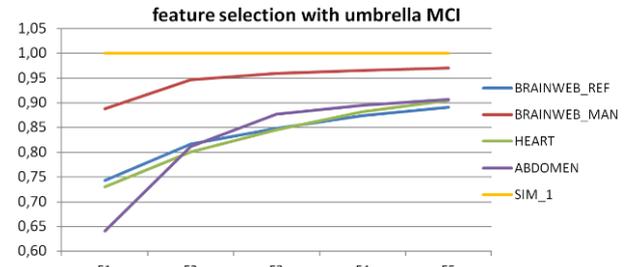
Figure 14: Overall P_{mcu} for increased number of selected *local* features.
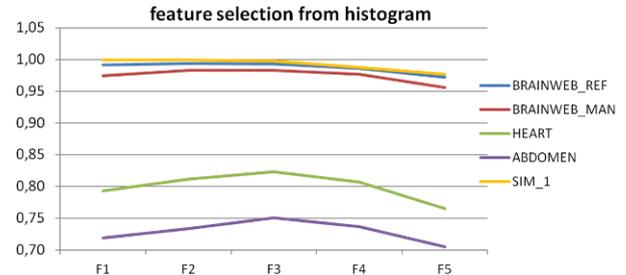
Figure 15: Overall P_{hist} for increased number of selected *local* features.

For the simulated datasets *SIM_2-SIM_8*, the results on automated local feature selection, as shown in Table 8, as well as automated feature selection of the best meta features, as presented in Table 9, accord with the previous findings on domain-specific single feature predictability as argued before.

Regarding the histogram-based overall predictability calculation strategy P_{hist}, density of the feature space with respect to the chosen dimension-dependent bin count as well as the applied filter parameterization has to be evaluated. Results are presented for *BRAINWEB_MAN* dataset in Fig. 16 by visualizing the 3D histogram for different feature sets. With features (F_5, F_{12}, F_{39}), due to outliers in the region size dimension, density of the feature space is low and

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

33

would necessitate for histogram equalization (a). Geometric positional features (F_{23}, F_{24}, F_{25}) allow refection of the class morphology itself (b). The feature set (F_1, F_8, F_{25}) that is considered best according to automated feature selection, shows a dense feature space and well-defined decision boundaries between the neighbouring class distributions (c).

Table 8: Selected *local* features for simulated datasets.

dataset	selected $\mathcal{F}_i$			predictability $P_{\mathcal{C}_{mcu}}$			
	$\mathcal{F}_I$	$\mathcal{F}_{II}$	$\mathcal{F}_{III}$	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$P_{total_{mcu}}$
SIM_2	41	24	23	1.00	.539	.536	.7232
SIM_3	41	25	24	1.00	.804	.809	.8840
SIM_4	5	21	33	1.00	.996	.996	.9974
SIM_5	41	42	43	1.00	.497	.772	.7805
SIM_6	40	41	43	1.00	1.00	1.00	1.000
SIM_7	3	23	24	1.00	.200	.842	.8412
SIM_8	42	39	8	1.00	.755	.883	.8916

Table 9: Selected *meta* features for simulated datasets.

dataset	selected $\hat{\mathcal{F}}_i$			predictability $P_{\mathcal{C}_{mcu}}$			
	$\hat{\mathcal{F}}_I$	$\hat{\mathcal{F}}_{II}$	$\hat{\mathcal{F}}_{III}$	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$P_{total_{mcu}}$
SIM_2	14	17	19	1.00	.892	.619	.8371
SIM_3	14	38	36	1.00	.983	.974	.9858
SIM_4	18	14	26	1.00	.947	.951	.9661
SIM_5	35	17	13	1.00	1.00	1.00	1.000
SIM_6	15	13	14	1.00	1.00	1.00	1.000
SIM_7	20	13	14	1.00	1.00	1.00	1.000
SIM_8	20	14	13	1.00	.875	.876	.9173

| (a) | (b) | (c) |
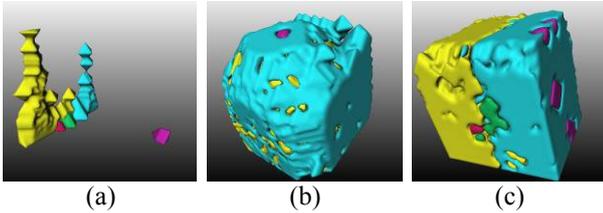
Figure 16: Visualization of 3D histogram with different utilized feature sets, showing decision boundaries for *white matter* (red), *grey matter* (green), *background* (magenta), *tissue* (yellow) and *remaining* voxel (cyan) from *BRAINWEB_MAN* datasets.

While feature selection can be parameterized to maximize region or voxel classification respectively, also predictability of the particular classes can be evaluated in detail, as shown in Fig. 17 for $k=13$ classes of *ABDOMEN* sequence. With features additionally incorporated for classification lead to higher overall predictability, also the predictability of the particular classes C_i increases by trend. Generally, for tomographic patient datasets, particular structures can only be classified at low confidence due to overlapping in the feature distributions. Thus, the incorporation of meta features is highly required to further improve distinctness for classification. Feature selection can also be adjusted to balance single class results or apply specific weights, besides tying to maximize region or voxel classification confidence.
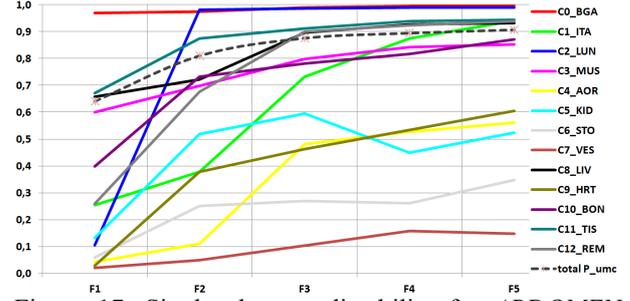
Figure 17: Single class predictability for *ABDOMEN* sequence according to increased number of features.

5.3. Results on Final Classification

Finally, classification of the pre-segmented datasets according to the automatically selected features has to be evaluated utilizing conventional classifiers from the machine learning domain. Segmentation results from neural network classification utilizing *HeuristicLab* (Wagner 2014) on *BRAINWEB_MAN* sequence are presented in Fig. 18. Besides some artefacts wrongly connected with grey and white matter volume and some marginal missing parts of the ventricle, the achievable classification results are of high quality.

For the three classes *grey matter*, *white matter* and *ventricle* a region classification precision of *.952* is achieved, incorporating *k=10* local features. With respect to the voxel error introduced by hybrid watershed pre-segmentation of *.969*, overall voxel classification precision is to be quantified with *.914*.

| (a) | (b) |
| (c) | (d) |
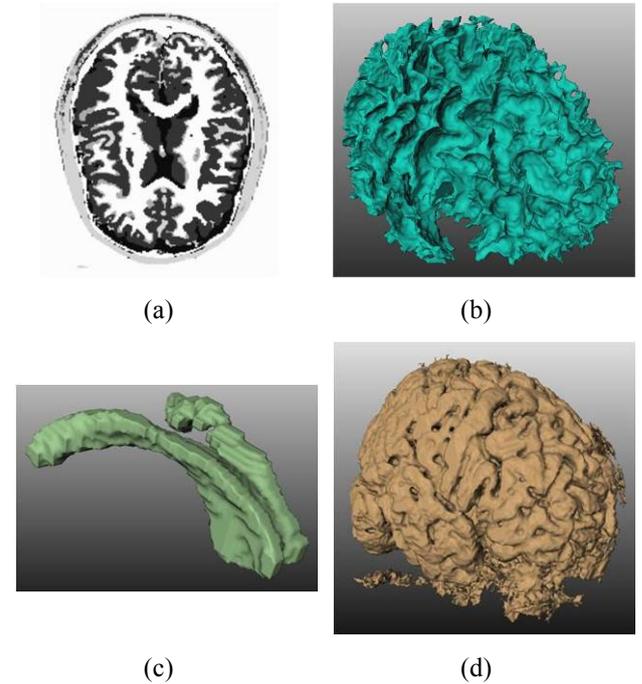
Figure 18: Classification results on *BRAINWEB_MAN* achieved with neural networks. Axial slice of labelled classes in (a) and surface renderings of *white matter* (b), *corpus callosum* (ventricle) (c) and the *grey matter* (d).

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

34

6. DISCUSSION AND CONCLUSIONS

A strategy for domain-specific selection of local features with respect to maximized region of voxel classification precision has been presented, also adjustable to different objectives by user-defined weights.

Introduced meta features provide potential for additional predictability in the medical segmentation and classification domain and are perfectly suited to be incorporated for genetic algorithms. Future developments will focus on application of these meta features to better steer the partially stochastic classification results of single individuals towards the optimal solution in evolutionary algorithms.

Besides, additional classifiers like genetic programming, support vectors or random forests will be evaluated with respect to achievable classification precision.

With respect to histogram-based approaches, sparseness of the feature space at higher dimensions will necessitate the future utilization of histogram equalization.

REFERENCES

Beyer, T., Schwenzer, N., Bisdas, S., Claussen C.D., and Pichler, B.J., 2010. MR/PET – Hybrid Imaging for the Next Decade. In *MAGNETOM Flash 3/2010*.

Boost, 2014. Boost C++ libraries, http://www.boost.org/, last visited 14/5/2014.

Boser, B.E., Guyon, I.M., Vapnik, V., 1992. A training algorithm for optimal margin classifiers. In *Proc. of Computational Learning Theory*, 144-152.

Cocosco, C.A., Kollokian, V., Kwan, K.-S., Evans, A.C., 1997. BrainWeb: Online Interface to a 3D MRI Simulated Database. In *NeuroImage* 5(4), 425.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1992. Training Models of Shape from Sets of Examples. In *Proceedings of the British Machine Vision Conference*, 9-18., Leeds, U.K.

Cootes, T.F., Edwards, G.J., Taylor, C.J., 1998. Active Appearance Models. In *Proceedings of the 5th European Conference on Computer Vision*, 484-498. June 2-6, Freiburg (Germany).

Felipe, J., Traina, A.J.M., Traina, C., 2003. Retrieval by Content of Medical Images Using Texture for Tissue Identification. In *Proc. of the 16th IEEE Symposium on Computer-Based Medical Systems CBMS'03*, 175-181.

Fürst, D., Schrempf, A., 2012. PATIENTSIM – Development of an Augmented Reality Simulator for Surgical Training of Vertebroplasty and Kyphoplasty. In *Proceedings of the 9th IASTED*.

Gonzalez, R.C., Wintz, P., 1987. *Digital Image Processing*. London: Addison Wesley.

Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, USA: Addison-Wesley Longman.

Kitware, Inc., 2014. ITK- The Insight Segmentation and Registration Toolkit, http://www.itk.org/ITK/, last visited 14/5/2014.

Koza, J.R., 1992. *Genetic Programming*. Massachusetts, USA: MIT Technology.

McInerney, T., Terzopoulos, D., 1996. Deformable Models in Medical Image Analysis : A Survey. In *Medical Image Analysis* 1 (2):91-108.

Metropolis, N., 1987. The Beginning of the Monte Carlo Method. In *Los Alamos Science*, 125-130.

Osher, S., Sethian, J.A., 1988. Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. In *Journal of Computational Physics* 79:12-49.

Schenk, A., Prause, G.P.M., 2001. Optimierte semi-automatische Segmentierung von 3D Objekten mit Live-Wire und Shape-Based-Interpolation. In *Bildverarbeitung für die Medizin* 2001, 202-206.

Swoboda, R., Backfrieder, W., Zwettler, G., Pfeifer, F., 2008. MIPP: A Medical Image Processing Platform based on ITK, VTK and Eclipse RCP. In *The Insight Journal* 3(2), 1-36.

Vapnik, V., 2000. *The Nature of Statistical Learning Theory*. New York, USA: Springer.

Vincent, L., Soille, P., 1991. Watersheds in Digital Space – An Efficient Algorithm Based on Immersion Simulations. In *IEEE Transactions on Pattern Analysis and Machine Intelligence 13(6)*, 583-598.

Wagner, S., 2014. Architecture and Design of the Heuristiclab Optimization Environment. In *Advanced methods and Applications in Computational Intelligence, Topics in Intelligent Engineering and Informatics Series*, Springer, 197-261.

Zwettler, G., Backfrieder, W., Swoboda, R., Pfeifer, F., 2009. Fast Fully-automated Model-driven Liver Segmentation Utilizing Slicewise Applied Levelsets on Large CT Data. In *Proceedings of the 21st European Modeling and Simulation Symposium*, 161-166.

Zwettler, G., Backfrieder, W., 2012. A New Hybrid Algorithm Based on Watershed Method, Confidence Connected Thresholding and Region Merging as Preprocessing for Statistical Classification of General Medical Images. In *Proceedings of the 24th European Modeling and Simulation Symposium EMSS2012*, 73-81.

Zwettler, G., Backfrieder, W., 2013. Generic Model-Based Application of Modular Image Processing Chains for Medical 3D Data Analysis in Clinical Research and Radiographer Training. In *Proc. of the International Workshop on Innovative Simulation for Healthcare IWISH2013*, 58-64.

Zwettler, G. and Backfrieder, W., 2014. Simulation of Tomographic Medical Image Data for Training of Generic Segmentation Models Utilizing Multivariate Feature Classification. In *Proc. of the SummerSim2014*, Monterey, USA, pp.10 (article in press).

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2014
978-88-97999-43-0; Bruzzone, Frascio, Longo, Merkuryev, Novak Eds.

35