This is an author post-print version of a published article: A. Kontunen et al., "Real Time Tissue Identification From Diathermy Smoke by Differential Mobility Spectrometry," in *IEEE Sensors Journal*, 2021, vol. 21, no. 1, pp. 717-724. Final form available online at https://doi.org/10.1109/JSEN.2020.3012965

Real Time Tissue Identification from Diathermy Smoke by Differential Mobility Spectrometry

Anton Kontunen, Markus Karjalainen, Anna Anttalainen, Osmo Anttalainen, Mikko Koskenranta, Antti Vehkaoja, Niku Oksala, and Antti Roine

Abstract

Current methods for intraoperative surgical margin assessment are inadequate in terms of diagnostic accuracy, ease-of-use, and speed of analysis. Molecular analysis of tissues could potentially overcome these issues. A system based on differential ion mobility spectrometry (DMS) analysis of surgical smoke has been proposed as one potential method, but to date, it has been able to function in a relatively slow and heavily controlled manner that is inadequate for clinical use. In this study, we present an integrated sensor system that can measure a surgical smoke sample in seconds and relay the information of the tissue type to the user in near real time in simulated surgical use. The system was validated by operating porcine adipose tissue and muscle tissue. The differentiation of these tissues based on their surgical smoke profile with a cross-validated linear discriminant analysis model produced a classification accuracy of 93.1% (N = 1059). The measurements were also classified with a convolutional neural network model, resulting in a classification accuracy of 93.2%. These results indicate that the DMS-based smoke analysis system is capable of rapid tissue identification from surgical smoke produced in freehand surgery.

1. Introduction

Surgical removal of the tumor mass is the mainstay of treatment in most solid cancers. Incomplete removal predisposes the patient to local recurrence of the tumor [1]. The goal is to remove the tumor with a margin of healthy tissue. Excessive removal of healthy tissue is associated with cosmetic or functional deficit and conservation of as much healthy tissue as possible is recommended. In the case of breast cancer, this is illustrated by the preference of *lumpectomy* (partial removal of the breast) over *mastectomy* (total removal of the breast) [2]. After the surgery, microscopic examination of the removed specimen is conducted to assess margin clearance. In 20–30% of cases,

cancerous tissue is found in or close to the tissue margin, often necessitating a reoperation, which negatively impacts patient's quality of life and yields extra costs [3]–[6]. This results in a growing need to create a tool for intraoperative tissue analysis that could reduce reoperations.

Currently, intraoperative margin monitoring is commonly conducted by imaging of the removed specimen, or by manual palpation and preoperative marking of the tumor using radiological guidance. However, certain subtypes cannot be identified by manual palpation. In these cases, microscopic analysis of frozen tissue sections or cytology smears have been shown to be effective but they are time consuming and expensive [7]–[9]. Optical coherence tomography (OCT) is a novel method that produces a microscopic view on the margin without preparation, but the interpretation of the output requires the expertise of a pathologist [10], [11].

Radiofrequency spectroscopy (RFS) detects cancerous tissue in the margin by assessing the dielectric properties of the specimen. The method has already been approved for clinical practice but has not gained wide acclaim due to poor specificity and difficulties orientating the positive finding, since the examination has to be performed *ex vivo* on removed tumor outside the patient [12].

Mass-spectrometry(MS)-based methods take tissue analysis from microscopic visual assessment to the analysis of the molecular composition of the tissues. Rapid evaporative ionization mass spectrometry (REIMS) is an experimental but promising technology that analyzes the surgical smoke that evaporates from the tissues that are cut during electrosurgical removal of the tumor [13], [14]. REIMS studies have demonstrated differences in the relative abundance of triglycerides and certain phospholipids, such as phosphatidylcholines and phosphatidylethanolamines between malignant and benign tissues. The proposed underlaying mechanism between these differences is the Warburg effect that has been well documented in various cancers [15]. REIMS is capable of differentiating benign and cancerous tissues at high accuracy within seconds during the surgery, thereby not interrupting the procedure or causing challenges in tumor orientation. Other MS-based methods include, SpiderMass, Picosecond InfraRed Laser Mass Spectrometry (PIRL-MS), and MasSpec pen [16]–[19]. SpiderMass and PIRL-MS use laser excitation to produce the gas phase sample, whereas in MasSpec pen, a water droplet is used to extract molecules from the target tissue sample. Research with these methods has been promising. Moreover, they are non-destructive by nature. However, the instruments disrupt the normal workflow of the surgeon by requiring an additional tool to be used during the surgery. All these technologies are hindered by their large physical size and the high cost of MS instrumentation.

In contrary to MS, our method is based on the usage of more affordable, almost maintenance free, differential mobility spectrometry (DMS) technology, in which gaseous compounds can be differentiated by their ion mobility characteristics in an asymmetrical electric field. Our previous research has demonstrated the potential of the DMS-based automatic tissue analysis system (ATAS) in the classification of healthy tissues, breast cancer and brain cancer [20]–[22]. In ATAS, the tissue sample is cut by controlling the electrosurgical knife with a modified 3D printer.

When the tissues are sampled with the ATAS, the electrosurgical cuts are standardized (depth, duration), thus stabilizing the smoke sample concentration [20]. However, as we intend to move towards the surgical application, the system needs to be able to function robustly with varying sampling conditions with close resemblance to a clinical situation and application. Therefore, ATAS needs to be replaced with handheld instrumentation. This in turn leads to technical requirements for the system that demand innovative engineering solutions and extensive modifications to the system compared to previously published methodology.

In this study, we introduce a tissue analysis system compliant with operation room workflow and standard

electrosurgical systems. We also describe its key engineering novelties, which include a heating solution to minimize contamination, an adaptive electric filter, a method to detect the induced diathermy current in order to time the measurement, and a surgical user interface. Furthermore, we evaluate the applicability of the system in the near real time discrimination of two porcine tissues by using machine learning methods and employ convolutional neural networks in the analysis of DMS data. The DMS-based sensor system and its technical advancements introduced in this study mark a potential breakthrough for a clinical application.

2. Materials and methods

a) Study material

The specimens used in this study were fresh commercial meat products that had distinctive areas of skeletal muscle tissue and adipose tissue as seen in Fig. 1. The use of human tissues in a proof-of-concept study is not justified due to the scarceness of the material and ethical aspects. Therefore, we chose to use porcine tissue samples (porcine flank) to demonstrate the function of the system. Another reason is that the tissues found in flank are operated in almost every electrosurgical procedure, due to them being in the way or around the target tissue, for example in cancer surgeries, which makes their identification clinically relevant.

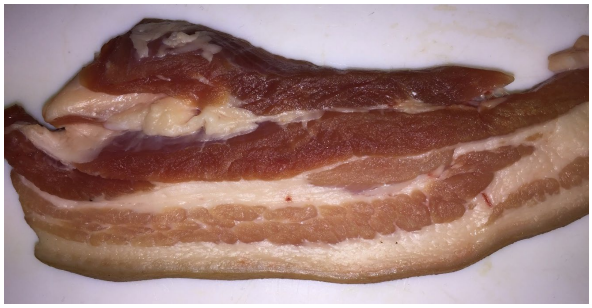

Figure 1. Porcine flank sample with macroscopically visible areas of skeletal muscle tissue and adipose tissue.

b) Measurement system and operation principle

The measurement system comprises an electrosurgical knife, surgical smoke evacuator, sample-preprocessing unit, and a DMS sensor. Fig. 2 shows a schematic representation of the measurement system with a pneumatic diagram of the sample-preprocessing unit. Fig 3. shows the measurement system in a laboratory setting.

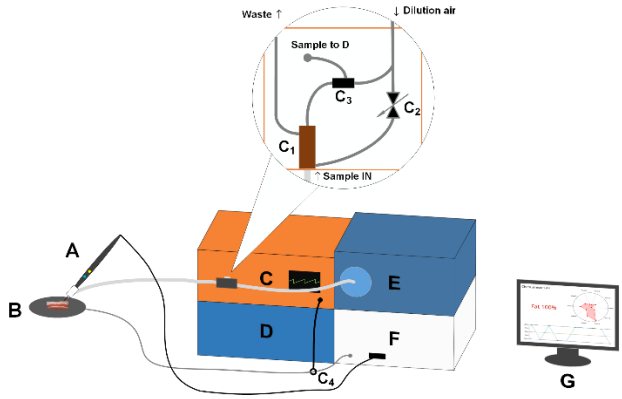

Figure 2. A stylized depiction of the measurement system with a pneumatic schematic of the sample pre-processing unit. A: Diathermy instrument with the attached 3-meter-long suction tube. B: Tissue specimen on top of the dispersive electrode. C: Sample pre-processing unit with the corona discharge filter (C_1), PID-controlled proportional valve (C_2), ejector dilution (C_3), and an inductive coil to detect the diathermy current (C_4). D: ENVI-AMC DMS sensor. E: Smoke evacuator. F: Diathermy unit. G: Surgical user interface.

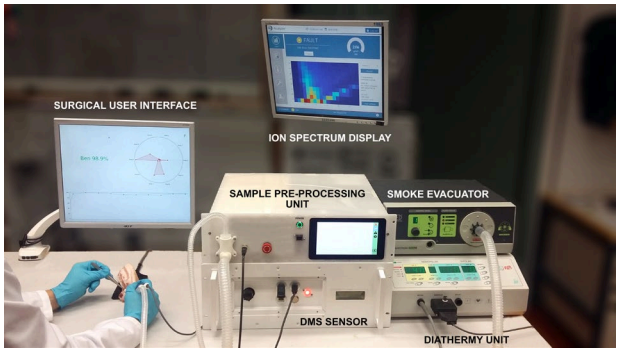

Figure 3. The measurement system in a laboratory setting.

In this study, the surgical cuts were generated by a monopolar knife electrode (A. in Fig. 2, HF 9805-24, HEBUmedical GmbH, Germany) attached to a diathermy unit (F. In Fig. 2, Itkacut 350MB, Innokas Medical Oy, Finland). The diathermy unit was set to operate at a nominal power of 40 W in 100% cut mode. The surgical smoke created during the cuts was collected by a three-meter-long suction tube (Handle aspiration kit, LED SpA, Italy) attached to the side of the diathermy instrument. From the side of the instrument, the smoke sample travelled to a T-junction that split the sample flow into two streams that led to a surgical smoke evacuator (E. in Fig. 2, SURTRON® EVAC, LED SpA, Italy) and the custom-built sample pre-processing unit (C. in Fig. 2).

The sample pre-processing unit of the current system has important additions and improvements compared to the previously described ATAS system [20], [21]. The main function of the pre-processing unit is to filter and dilute the smoke sample. However, previous studies have shown that the carry-over signal from preceding measurements is a problem when conducting rapid consecutive measurements. Therefore, the sample line in the pre-processing unit of the improved system was minimized to 75 cm to decrease the available scent adsorption area. In addition, the sample line was heated by a resistive cable heater and enclosed in a wooden box made from 4 mm thick birch plywood that keeps the sample pneumatics at elevated temperature (70°C), allowing for shorter recovery time due to reduced condensation

and scent adherence, and increased desorption of the scent molecules from the tubing. In theory, a higher temperature would work even better in mitigating the carry-over phenomenon, but 70°C was chosen as the limit to ensure that the electronic components within the pre-processing unit do not malfunction due to excessive heating.

Besides the integrated heating and minimized sample travel length, the main technical novelty of the pre-processing unit is an improved dilution system that controls the smoke sample intake based on the concentration. The adaptive dilution system is built around a filtration solution (C_1 in Fig. 2) that utilizes a corona discharge to remove particulate matter from the sample stream. A schematic representation of the filter is presented in Fig. 4. The discharge is created by applying a 5 kV voltage to the center electrode, while keeping the copper casing at ground potential. This ionizes the smoke sample and forces the particulate matter to the casing, from where it continues to a waste stream.

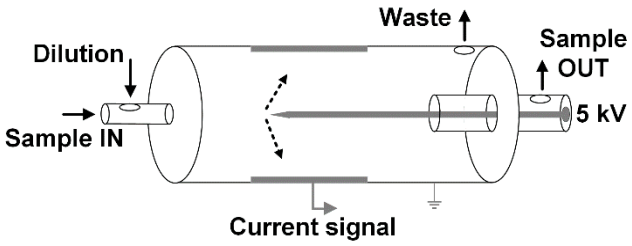

Figure 4. Corona discharge particulate filter.

In addition to removing large particles from the sample stream, the new filter measures the amount of the particulate matter. The inner wall of the filter contains electrodes, which produce a current signal, when the ionized particulate matter reaches the surface. The smoke concentration depends on how long or deep the electrosurgical cut is. The sample is mixed with dilution air. Diluted concentration is regulated with a proportional-integral-derivative (PID) feedback-controlled valve (C_2 in Fig. 2). PID is a widely used control method, where a deviation of a desired setpoint is corrected by using three corrective terms in unison to achieve minimum overshoot and delay in reaching the desired output values. In terms of the real time tissue analysis system, this means that when particulate matter causes a current signal deviation in the filter, the PID-controlled valve tries to actively compensate the change by increasing the amount of dilution air without overdiluting the sample. We have previously studied that even with similar cutting parameters, the amount of particulate matter is highly dependent on the tissue type [23]. This means that an adaptive dilution system is a useful solution to ensure that the DMS measurement does not saturate regardless of the tissue type or cutting conditions.

In addition to the adaptive dilution system, the sample pre-processing unit has a Raspberry Pi 7” Touch Screen Display (Raspberry Pi Foundation, UK) that relays information to the user about the electrical current signal of both the corona filter and the diathermy knife. The signal from the knife is monitored by an induction coil (C_4 in Fig. 2) wrapped around the cable of the dispersive electrode (B. in Fig. 2). The monitoring of the induced cutting current was utilized in the synchronization of the DMS measurement with the smoke production. In other words, the DMS measurement was triggered, when the current of the electrosurgical knife was above a fixed threshold that was exceeded only upon contact with tissue material.

The DMS sensor used in the study is an Envi-AMC® spectrometer (D. in Fig. 2) manufactured by Environics Oy, Finland. The sensor is capable of parts-per-billion-level detection of gaseous substances and operates with separation

fields between 10 to 80 Townsend (Td) with <0.01 Td resolution. The waveform is created with a field-effect transistor based half-bridge with 250 kHz operating frequency resulting in a square waveform. The waveform is run with 5 % duty cycle. The compensation field can be varied between ±1.5 Td with resolution <0.001Td. The DMS filter is arranged from three planar plates forming two identical channels, both with channel height of 0.25 mm, width of 6 mm and length of 16 mm [24]. 2.8 liters per minute sample flow rate is created with a venture nozzle, driven by clean pressurized air. Sample is drawn untreated into the DMS filter via Am-241, 5.92 MBq ionization chamber. The residence time in the ionization area is about 16 ms and in the DMS filter about 1 ms. Filtered ions are collected with three sequentially arranged electrodes. The collection field is adjusted to collect most of the ions with the middlemost and the largest of the collection electrodes. The measurement is controlled by a LUA-language script, which allows detailed control for each of the measurement parameters such as flow, frequency, duty cycle, separation field settings, compensation field settings, signal noise filtering and selection of ion polarity. The operation principle of the DMS sensor is illustrated in Fig. 5.

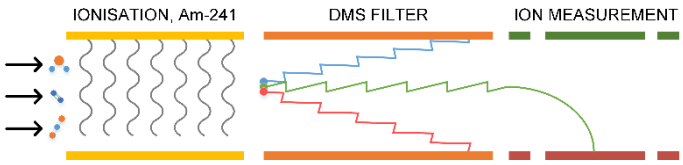

Figure 5. The operation principle of the ENVI-AMC® spectrometer. The black arrows represent the incoming gas sample molecules, and the colored lines represent sample ions that diverge in the DMS filter based on their differential mobility characteristics.

The result of a DMS measurement can be represented as a colored 2D image (i.e. dispersion plot), where the color represents the number of ions that reach the detector with different DMS filter separation and compensation voltage values. All DMS data in this paper were measured with the following settings: Positive ion mode was used, the compensation voltage sweep (x-axis in the dispersion plot) was set from -0.8 V to 5 V with 25 steps, and the separation voltage sweep (y-axis in the dispersion plot) was set from 340 V to 740 V with 8 steps. With these settings, the measurement duration was approximately 5 seconds.

After the DMS measurement of a smoke sample is finished, the data is stored to a Microsoft® Azure (Microsoft, U.S.A) cloud storage for possible offline analysis. The measurement data is also transferred to a Raspberry Pi computer (Model 3B). The computer has a linear discriminant analysis (LDA) classification model that is used to classify the newly measured data. LDA has been used as the primary classification method in our previous studies with good results and it has also been used, for example, in the REIMS studies [20], [21]. LDA classification is computationally simple and the result is produced in milliseconds. The classification script of the real time system was written in Python.

The Raspberry Pi computer that is used for online classification also displays the result in an attached monitor that functions as the surgical user interface (UI). The surgical UI (G. in Fig. 2) can be set to display the result of the LDA classification as the predicted tissue type alongside the posterior probability of the prediction. In other words, the UI can tell the surgeon how certain the model is of its prediction. This enables a more nuanced assessment of the operated tissue than a completely binary statement. In addition, the results of the previous classifications can be displayed with a graph, making following the analysis much easier.

The whole process from the tissue evaporation by freehand surgery to displaying the classification result in the surgical UI takes approximately 9 seconds. After the result is displayed, the process can start again, when the trigger level for the diathermy current exceeds the threshold level (i.e. the knife is in contact with the tissue). Fig. 6 shows the operation principle of the system in the form of a flowchart.

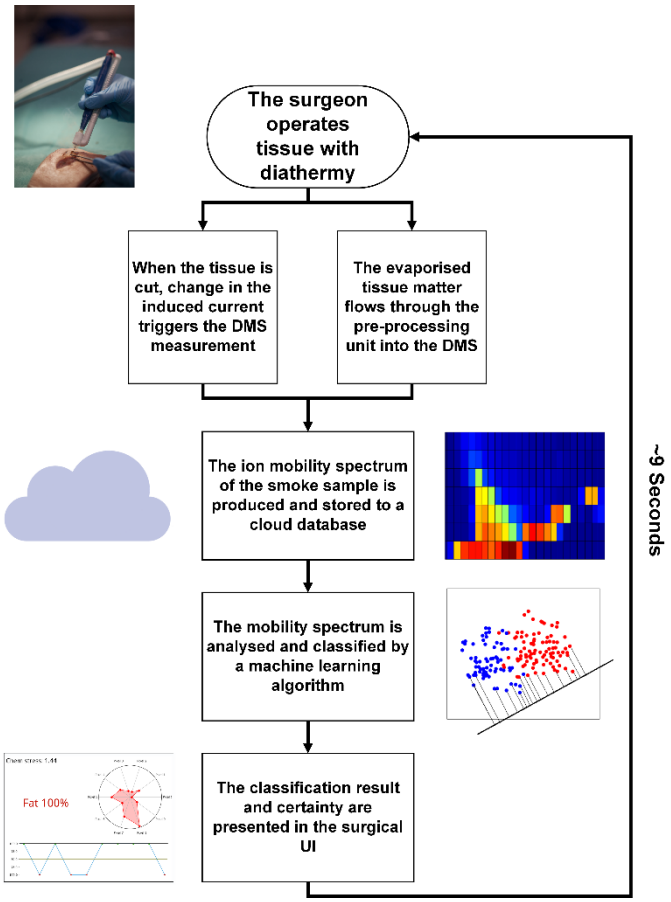

Figure 6. Flowchart of the operation principle of the real time tissue analysis system.

c) Measurement protocol and tissue sampling

The measurement protocol was designed to resemble the use of diathermy in surgical practice as closely as possible. An experienced consultant surgeon (N. Oksala) participated in the design and implementation of the protocol. The protocol consisted of three steps: 1. A fresh tissue specimen was placed on the dispersive electrode of the diathermy unit; 2. The tissue specimen was incised for 1–3 seconds with the diathermy instrument. Each incision triggered a DMS measurement of the surgical smoke. 3. A recovery period of approximately 10 seconds was taken to reduce carry-over.

The process steps 2. and 3. were repeated until no uncut surface area was available. This equated to approximately 150-200 incisions per specimen. At the end of the measurement session, the specimen was disposed, and the dispersive electrode was cleaned with isopropanol.

In total, 1159 freehand smoke samples (580 to skeletal muscle and 579 to adipose tissue) from seven specimens were measured. Out of the 1159 samples, 1059 were used for classification model training and internal validation and 100 were used for external validation. To reduce operator-dependent and temporal bias, the sample cuts were made by three operators over a period of seven measurement days.

d) Data analysis and classification

The features of the DMS data can be presented as color-coded pixels in the two-dimensional dispersion plot, where the color represents the 16-bit signal intensity value. In this respect, the DMS output dispersion plots can be considered as images that depict the chemical fingerprint of a substance, which means that image recognition methods are a viable option for the analysis and classification of DMS data. Thus, in addition to the simple LDA classification, we tested convolutional neural networks (CNN) for the classification of the tissue smoke data.

Data produced with the system were analyzed offline with MATLAB (Version R2017b, The MathWorks Inc., U.S.A). Before classification, the values of the DMS data were offset-corrected to have only positive values and normalized between 0 to 1. The offline models were implemented to have more numerical data of the performance of the classification models than what was used in the real time classification. The CNNs were created using the mxnet package for R (R Core Team, 2018). The CNNs consisted of two convolutional and two fully connected layers. The architectural and regularization parameters were chosen based on a grid search and heuristic testing. The same data was used in the hyper-parameter selection and is reported as 10-fold cross-validated (10-f-CV) results, which may have caused some positive bias in the results. To our knowledge, there are no previous publications, where CNNs have been used for DMS data.

For the performance evaluation with the training data, the models were 10-fold cross validated. In addition to the internal cross-validation, the LDA classification model was validated with an independent data set that was measured four months after the training data. The external validation with the independent test set was deemed necessary in order to assess the effect of long-term variation in the produced DMS outputs on classification performance. The classification results of the test set were analyzed both with the Raspberry embedded real time model, and offline after the sampling. The test set consisted of 100 cuts (50 skeletal muscle and 50 adipose tissue).

Further analysis of the DMS data and its features for classification were also explored by utilizing a process called forward sequential feature selection (FSFS). FSFS is an iterative process that aims to find the most significant features for the differentiation of the classes by selecting the features one-by-one until the classification performance does not significantly improve [25]. The feature selection for the LDA classification model was repeated 1000 times with different 10-fold cross-validation partitions to ensure that the selected features would not be dependent on the training set.

3. Results

The data analysis from the metadata of the DMS measurements revealed that the median time between cuts with the freehand sampling was 18 seconds (median absolute deviation 4.4 seconds). The median time is twice as long as the

minimum time for consecutive measurements, since the point of the training data was to get good representative samples of both tissues, rather than trying to minimize the analysis time.

In internal validation, the mean classification accuracy of the 10-f-CV LDA model was 93.1%, with sensitivity of 91.5% and specificity of 94.7%. The CNN model produced a mean cross-validated classification accuracy of 93.2%, with sensitivity of 91.9% and specificity of 94.5%. The results and the internal validation are presented in Table I as confusion matrices, where the true class is represented in the rows and predicted class in the columns.

In terms of external validation with the independent set, the accuracy of the real time classification was 87.0%, with sensitivity being 80.0% and specificity being 94.0%. This means that out of the 100 test samples 13 were misclassified by the real time Python model during freehand sampling. A confusion matrix of the real time classification along with the associated offline CNN classification is presented in Table II. The classification of the independent data set with the CNN model produced a 50.0% classification accuracy with all predictions being muscle tissue, i.e. sensitivity 100.0% and specificity 0%.

Due to the guess level predictions produced by the CNN model in the external validation, the training set of the model was increased with 1240 samples (721 samples of skeletal muscle and 519 samples of adipose tissue) that were not measured by freehand surgery, but instead with the ATAS system in a different study. This increased the robustness of the model and resulted in a classification accuracy of 88.0%, with sensitivity of 100.0%, and specificity of 76.0%, with the independent test set. When the additional data was also used for the training of the LDA model, the classification accuracy in offline analysis increased to 96.0%, with sensitivity of 100.0% and specificity of 92.0%. Table III shows the confusion matrices of the external validation with the CNN and LDA models trained with additional data.

The 1000 FSFS cycles with the original 1059 sample data set yielded a median number of 10 selected features from the full spectrum. By using these features for classification of the tissue types, the cross-validated mean classification accuracy was 93.2%, with sensitivity of 95.9% and specificity of 90.4%. The ten selected features that explain the differences in the spectra of muscle tissue and adipose tissue the most, are presented in Fig. 7, along with example outputs produced by both tissues.

Table 1.

Classification results with 10-fold cross validation. The true class is in rows and the predicted class in columns

	LDA			CNN	
	Fat	**Muscle**		**Fat**	**Muscle**
Fat	501	28	**Fat**	500	29
Muscle	45	485	**Muscle**	43	487

Table 2.

Classification results with the independent external test set. The true class is in rows and the predicted class in columns

	LDA			CNN	
	Fat	**Muscle**		**Fat**	**Muscle**
Fat	47	3	**Fat**	0	50
Muscle	10	40	**Muscle**	0	50

Table 3.

Classification results with the independent external test set with models that were trained with additional data. True class is in rows and predicted class in columns

	LDA			CNN	
	Fat	Muscle		Fat	Muscle
Fat	46	4	Fat	38	12
Muscle	0	50	Muscle	0	50

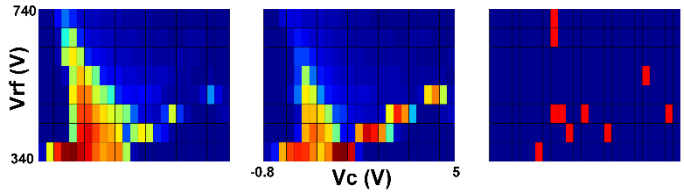

Figure 7. Example dispersion plot of muscle tissue (left), adipose tissue (center), and ten features selected in the FSFS process for the classification of the tissues (right). Vrf is the DMS separation voltage and Vc is the DMS compensation voltage.

4. Discussion

a) Findings and impact

Our results show that the latency, classification accuracy and reliability of our freehand-operated tissue analysis system are acceptable and pave way towards clinical validation of the technology. The study works as a proof-of-concept for previously unexplored application of rapid DMS-based tissue identification. However, some limitations that hinder the integration of the system to a clinical setting still remain to be resolved.

b) Limitations

As this study focused on describing the technical novelty of the improved surgical system, we cannot emphasize the clinical relevance of the identification results and their impact. However, the simplified binary setting works as a promising surrogate for a real surgical situation since the intended application of the system would be to differentiate malignant and benign tissues in a similar manner. In addition, muscle and adipose tissue are usually the most prominent types of healthy tissue surrounding tumors, which makes their identification relevant.

In the intended final surgical application, one of the most desired attributes of the system is minimal analysis time, so that the surgeon can perform the operation without disturbances in the workflow. The current minimum interval between the DMS measurements, approximately 9 seconds, should be improved in order to make the assessment of the resection

line easier. With this respect, our system is still inferior to the current MS-based solutions. With REIMS, the time between the start of the cut to displaying the result is reported to be less than two seconds [13]. The limiting factor for the measurement duration in our system is the ENVI-AMC DMS. As most DMS devices are not designed to perform rapid measurements, but rather to monitor ambient air quality, the measurements require a full sweep of voltage values and cannot be concentrated on some key areas of the spectrum. Currently, the only way to decrease the measurement time is to decrease the resolution or the spectrum window size, which will decrease the classification performance.

Even in the task of a rather simple classification between skeletal muscle and intermuscular fat, the system failed to achieve 100% accuracy. This presents a clear area for improvement regarding the diagnostic properties of the system. However, the main reason for the misclassification can be traced back to the variability on the external measurement conditions that affect the DMS output. Especially the changes in environmental humidity and temperature can cause differences between the spectra in terms of offsets, signal intensities, and even feature positions. This is apparent in the CNN results of the independent data set, which were at guess level. To account for the variability of the measured outputs over time, a calibration protocol with a known control substance should be created in future studies.

The reference annotation of the tissue specimens was conducted by macroscopic examination and it is possible that a subset of the muscle specimens contained significant amount of fat that may impair the results. However, the effect is likely small, and the risk of overoptimistic results is low.

In addition, due to a physical blockage in the airways of the DMS device that occurred between the measurements of the training data and the independent test set, the DMS core circuit boards were disassembled, cleaned, and reassembled before measuring the independent test set. The reassembly might have caused some change in the measurement output.

c) Analysis of the Results

Besides pointing out the significant negative effect of long-term output drift of the DMS device to classification accuracy, the CNN and LDA results showed that the performance of the classification model is largely dependent on the size and variability of the training set. Even with data that were not produced by freehand surgery, the classification performance for the independent data set increased with both models, when the size of the training set was increased. Furthermore, the CNN model was particularly affected by the added training data, since the classification accuracy increased from guess level prediction to 88%, i.e. the number of misclassifications was decreased by 76%. This is perhaps expected, since neural networks are highly dependent on the amount of training data and they are most effective in applications, where the size of the training set can be millions of samples [26]. However, these results present interesting research questions regarding the optimal data analysis of this type of DMS data that should be answered in future studies.

Another aspect that needs further exploration is the effect of carry-over signal to classification performance. While the integrated heating of the sample pre-processing unit and PID-controlled dilution systems were implemented and reduced the overall recovery time of the system, residual scent of previous measurements can accumulate and slowly release from the altogether 3-meter-long tubing before the pre-processing unit. The carry-over signal might explain part of the misclassifications in this study. However, in similar applications based on MS, the study groups have been able to get near perfect classification results with rapid measurements [14], [19], indicating that the possible carry-over signal can be distinguished from the relevant information of a new sample, at least in MS-based solutions.

The relevant information of the samples can be considered the key features that enable the differentiation based on the DMS spectra. However, some areas of the spectra do not yield any significant information regardless of the measured substance, but contain electronic noise derived from the DMS hardware. In other words, the spectra contain changes that are not related to the tissue type and thus including them as training data for the classifier can predispose the model to overfitting and reduce the generalizability. Much like in our previous study with breast cancer samples, the FSFS results revealed that the acquisition of the full DMS spectrum from each smoke sample is redundant in terms of classification accuracy [21]. In this study, by selecting only 5% (10/200) of the features, the classification accuracy increased by one decimal compared to the classification with the full spectrum. This means that if the DMS device would allow for selection of specific waveform voltage and compensation voltage values instead of full sweeps, the measurement time could be significantly shortened, while retaining similar, or even better, performance. This, in turn, would lead to better applicability of the system to the intended real time surgical use.

Since the aim of the system is qualitative assessment of tissue type and extensive libraries for DMS are not available, the direct identification of the measured molecules is not required. However, the features selected by the FSFS are indicative of the cellular content differences between muscle and fat. The water content between the tissues is different and this can be seen in the selected features in the rightmost selected features that are in line with the so-called reactant ion peak. The selected features on the left side of the spectra derive from differences in heavier molecules such as lipids and their degradation products.

5. Conclusion

We have shown that fast tissue identification during freehand diathermy cutting based on diathermy smoke analysis by DMS is possible. In addition to the previously used LDA classification, we employed a CNN model that has potential to improve with the addition of new data. In terms of diagnostic accuracy and overall performance, the DMS-based system is not yet capable of producing results that would compete with the MS-based systems, but its simplicity and robustness partly compensate its limitations. The difficulties associated with freehand smoke sampling have been mostly overcome with engineering solutions and the system will continue to improve with further research. In future studies, the performance of the system will be tested in a surgical setting with clinically relevant human tissues and the following results will determine the feasibility of the technology in clinical use.

References

[1] A. Bodilsen *et al.*, “The Influence of Repeat Surgery and Residual Disease on Recurrence After Breast-Conserving Surgery: A Danish Breast Cancer Cooperative Group Study,” *Ann. Surg. Oncol.*, vol. 22, no. 3, pp. 476–485, 2015.

[2] T. P. Ahern, H. Larsson, J. P. Garne, D. P. Cronin-Fenton, H. T. Sørensen, and T. L. Lash, “Trends in breast-conserving surgery in Denmark, 1982–2002,” *Eur. J. Epidemiol.*, vol. 23, no. 2, pp. 109–114, 2008.

[3] R. Jeevan *et al.*, “Reoperation rates after breast conserving surgery for breast cancer among women in England: Retrospective study of hospital episode statistics,” *BMJ*, vol. 345, no. 7869, pp. 1–9, 2012.

[4] S. Jendrian, K. Steffens, B. Schmalfeldt, E. Laakmann, C. Bergelt, and I. Witzel, "Quality of life in patients with recurrent breast cancer after second breast-conserving therapy in comparison with mastectomy: the German experience," *Breast Cancer Res. Treat.*, vol. 163, no. 3, pp. 517–526, 2017.

[5] Y. Grant *et al.*, "Patient-level costs in margin re-excision for breast-conserving surgery," *BJS*, vol. 106, no. 4, pp. 384–394, Mar. 2019.

[6] R. E. Pataky and C. R. Baliski, "Reoperation costs in attempted breast-conserving surgery: A decision analysis," *Curr. Oncol.*, vol. 23, no. 5, pp. 314–321, 2016.

[7] J. C. Boughey *et al.*, "Impact of analysis of frozen-section margin on reoperation rates in women undergoing lumpectomy for breast cancer: Evaluation of the National Surgical Quality Improvement Program data," *Surg. (United States)*, vol. 156, no. 1, pp. 190–197, 2014.

[8] J. M. Jorns *et al.*, "Intraoperative frozen section analysis of margins in breast conserving surgery significantly decreases reoperative rates: One-year experience at an ambulatory surgical center," *Am. J. Clin. Pathol.*, vol. 138, no. 5, pp. 657–669, 2012.

[9] E. R. St John *et al.*, "Diagnostic accuracy of intraoperative techniques for margin assessment in breast cancer surgery a meta-analysis," *Ann. Surg.*, vol. 265, no. 2, pp. 300–310, 2017.

[10] S. J. Erickson-Bhatt, R. M. Nolan, N. D. Shemonski, S. G. Adie, J. Putney, and D. Darga, "Real-time imaging of the resection bed using a handheld probe to reduce incidence of microscopic positive margins in cancer surgery," *Cancer Res*, vol. 75, no. 18, pp. 3706–3712, 2015.

[11] S. A. Boppart *et al.*, "Label-free optical imaging technologies for rapid translation and use during intraoperative surgical and tumor margin assessment," *J. Biomed. Opt.*, vol. 23, no. 2, pp. 1–10, 2018.

[12] F. Schnabel *et al.*, "A randomized prospective study of lumpectomy margin assessment with use of marginprobe in patients with nonpalpable breast malignancies," *Ann. Surg. Oncol.*, vol. 21, no. 5, pp. 1589–1595, 2014.

[13] E. R. St John *et al.*, "Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: Towards an intelligent knife for breast cancer surgery," *Breast Cancer Res.*, vol. 19, no. 59, pp. 1–14, 2017.

[14] D. L. Phelps *et al.*, "The surgical intelligent knife distinguishes normal, borderline and malignant gynaecological tissues using rapid evaporative ionisation mass spectrometry (REIMS)," *Br. J. Cancer*, vol. 118, no. 10, pp. 1349–1358, 2018.

[15] K. O. Alfarouk and C. T. S. and L. Schwartz, "The Warburg Effect and the Hallmarks of Cancer," *Anti-Cancer Agents in Medicinal Chemistry*, vol. 17, no. 2. pp. 164–170, 2017.

[16] B. Fatou *et al.*, "In vivo Real-Time Mass Spectrometry for Guided Surgery Application," *Sci. Rep.*, vol. 6, no. 25919, pp. 1–14, May 2016.

[17] M. Woolman *et al.*, "Rapid determination of medulloblastoma subgroup affiliation with mass spectrometry using a handheld picosecond infrared laser desorption probe," *Chem. Sci.*, vol. 8, no. 9, pp. 6508–6519, 2017.

[18] J. Zhang *et al.*, "Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld mass spectrometry system," *Sci. Transl. Med.*, vol. 9, no. 406, pp. 1–11, Sep. 2017.

[19] P. Saudemont *et al.*, "Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology," *Cancer Cell*, vol. 34, no. 5, pp. 840–851, 2018.

[20] A. Kontunen, M. Karjalainen, J. Lekkala, A. Roine, and N. Oksala, "Tissue Identification in a Porcine Model by

Differential Ion Mobility Spectrometry Analysis of Surgical Smoke," *Ann. Biomed. Eng.*, vol. 46, no. 8, pp. 1091–1100, 2018.

[21] M. Sutinen *et al.*, "Identification of breast tumors from diathermy smoke by differential ion mobility spectrometry," *Eur. J. Surg. Oncol.*, vol. 45, no. 2, pp. 141–146, Feb. 2019.

[22] I. Haapala *et al.*, "Identifying brain tumors by differential mobility spectrometry analysis of diathermy smoke," *J. Neurosurg. JNS*, vol. 1, no. AOP, pp. 1–7, 2019.

[23] M. Karjalainen *et al.*, "The characterization of surgical smoke from various tissues and its implications for occupational safety," *PLoS One*, vol. 13, no. 4, pp. 1–13, Apr. 2018.

[24] O. Anttalainen, J. Puton, K. Peräkorpi, E. Budzyńska, G. Eiceman, and M. Sillanpää, "Differential mobility spectrometers with tuneable separation voltage – Theoretical models and experimental findings," *TrAC Trends Anal. Chem.*, vol. 105, no. 8, pp. 413–423, 2018.

[25] S. Wang, J. Tang, and H. Liu, "Feature Selection BT," in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, pp. 503–511.

[26] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, no. 5, pp. 354–377, 2018.