# 9 Supplementary for

# Thyroid Cancer Detection and Classification Using Spectral Imaging and Artificial Intelligence

Maya Almagor[1,*], Yotam Shapira[1,*], Adam Soker[1], Gabor Fischer[2], John Gartner[2], Sabine Mai[3], Yuval Garini[1,4]

[1] Biomedical Engineering Faculty & Russell Berrie Nanotechnology Institute, Technion − Israel Institute of Technology, Haifa, Israel

[2] Department of Pathology, Max Rady College of Medicine, University of Manitoba, Winnipeg, Canada

[3] Department of Physiology and Pathophysiology, University of Manitoba, Winnipeg, Canada

[4] Russell Berrie Nanotechnology Institute, Technion − Israel Institute of Technology, Haifa, Israel

* Equal contribution

Corresponding author details: Yuval Garini - yuval.garini@bm.technion.ac.il

Yotam Shapira – yotamshapira@campus.technion.ac.il

## 9.1 Feature-Based Classification

The random forest classifier utilizes a set of features extracted from the segmentation masks that automatically detect the nuclei within an input image. These features, which are either morphological or spectral in nature, serve as input to the classifier to assign each nucleus to one of two categories: normal or cancerous. The selection of these features is based on their relevance for training the model and their applicability to new cases.

- **Nucleus size:** the total number of pixels that make up the nucleus in the image. Tumor nuclei are typically larger than normal nuclei.
- **Average nucleus spectrum:** The mean intensity of the nucleus at different wavelengths, as measured by the SI system.
- **Spectral variance across the nucleus:** The variance of the spectral intensities at different wavelengths within the nucleus. Tumor nuclei tend to exhibit more heterogeneous spectra, with greater interchromatin space compared to normal nuclei.
- **Maximum measured intensity:** Tumor nuclei often display higher intensity in spectral images due to the different organization of the chromatin within the nucleus.
- **Spectrum normalization factor:** The difference maximum and minimum measured intensities within the nucleus. This factor normalizes the spectrum, emphasizing differences in its shape rather than its intensity.
- **Nucleus circularity:** A measure of the nucleus's roundness, calculated using the formula:

$$Circularity = \frac{4\pi \times (nucleus\ area)}{(nucleus\ perimeter)^2}$$

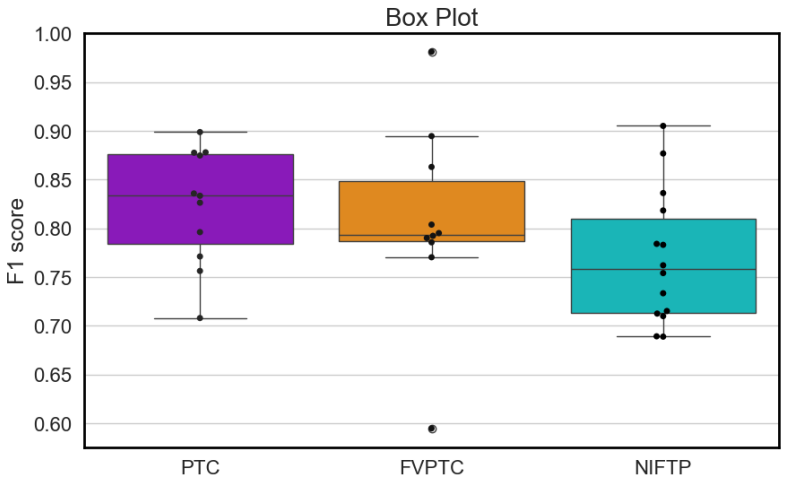
- **Number of neighboring nuclei within a defined radius:** we defined a radius of 50 µm around each nucleus. Cancerous cells tend to be more densely clustered than normal cells.
- **Standard deviation of neighboring nuclei:** The standard deviation of the number of neighboring nuclei within the specified radius, representing the variability across the entire region of interest, rather than focusing on a specific sub-region.

## 9.2 K-Means Results Summary

**Supplementary table 1**: Full results of k-means, per case, with and without nuclei exhibiting mixed characteristics.

| | **Including** mixed characteristics | | | | | **Excluding** mixed characteristics | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Case** | **TN** | **TP** | **F1** | **# normal** | **# cancer** | **TN** | **TP** | **F1** | **# normal** | **# cancer** |
| 18S_15062_NIFTP | 1554 | 3659 | 0.762 | 2186 | 5311 | 1435 | 3338 | 0.783 | 1937 | 4686 |
| 18S_44928_NIFTP | 1396 | 4869 | 0.877 | 1607 | 6025 | 1103 | 3638 | 0.860 | 1310 | 4610 |
| HS19_12314_NIFTP | 1438 | 2521 | 0.710 | 2228 | 3791 | 1335 | 2345 | 0.873 | 1534 | 2826 |
| HS19_18274_NIFTP | 1756 | 2381 | 0.733 | 2539 | 3329 | 1637 | 2186 | 0.742 | 2256 | 3085 |
| HS20_13920_NIFTP | 3227 | 4448 | 0.819 | 4003 | 5646 | 2996 | 4365 | 0.852 | 3647 | 5231 |
| HS20_21416_NIFTP | 1175 | 1812 | 0.713 | 1679 | 2770 | 1088 | 1517 | 0.686 | 1555 | 2436 |
| SS18_1993_NIFTP | 2986 | 3026 | 0.836 | 3627 | 3570 | 2678 | 2942 | 0.863 | 3243 | 3312 |
| SS18_199_NIFTP | 1028 | 1168 | 0.783 | 1346 | 1497 | 776 | 1086 | 0.789 | 1077 | 1367 |
| SS18_5274_NIFTP | 1983 | 2085 | 0.784 | 2570 | 2646 | 1855 | 1946 | 0.823 | 2247 | 2391 |
| SS19_12451_NIFTP | 406 | 3686 | 0.905 | 433 | 4431 | 356 | 3606 | 0.945 | 384 | 3998 |
| SS19_2344_NIFTP | 1819 | 2265 | 0.754 | 2471 | 3090 | 1660 | 2165 | 0.789 | 2170 | 2813 |
| SS19_4159_NIFTP | 1871 | 2801 | 0.715 | 2830 | 4073 | 1729 | 2605 | 0.728 | 2556 | 3723 |
| SS19_5490_NIFTP | 1679 | 2006 | 0.689 | 2380 | 3114 | 1500 | 1878 | 0.714 | 2052 | 2827 |
| SS19_8215_NIFTP | 1470 | 1437 | 0.689 | 2215 | 1990 | 1212 | 1242 | 0.696 | 1806 | 1732 |
| 18_S009289_PTC | 1316 | 3563 | 0.756 | 1848 | 5327 | 1107 | 3346 | 0.803 | 1490 | 4601 |
| 18_S013537_PTC | 1485 | 2680 | 0.708 | 1509 | 4867 | 993 | 2571 | 0.875 | 1028 | 3270 |
| 18_S013539_PTC | 1324 | 3831 | 0.875 | 1560 | 4691 | 1176 | 3760 | 0.912 | 1305 | 4353 |
| 18_S015988_PTC | 1384 | 3258 | 0.878 | 1598 | 3952 | 1256 | 3196 | 0.916 | 1385 | 3650 |
| 18_S020329_PTC | 2363 | 2949 | 0.796 | 3035 | 3787 | 2143 | 2819 | 0.822 | 2639 | 3537 |
| SS18_4375_PTC | 1010 | 1084 | 0.878 | 1134 | 1261 | 967 | 1063 | 0.956 | 985 | 1143 |
| SS18_4748_PTC | 250 | 1207 | 0.836 | 317 | 1614 | 219 | 1175 | 0.891 | 272 | 1409 |
| SS18_5367_PTC | 1506 | 1087 | 0.771 | 1857 | 1381 | 1290 | 1016 | 0.811 | 1530 | 1250 |
| SS19_12483_PTC | 3392 | 4724 | 0.899 | 3960 | 5219 | 2930 | 3979 | 0.905 | 3396 | 4344 |
| SS19_438_PTC | 1275 | 2777 | 0.826 | 1628 | 3591 | 1071 | 2295 | 0.816 | 1389 | 3013 |
| SS19_785_PTC | 632 | 811 | 0.834 | 770 | 997 | 344 | 365 | 0.991 | 351 | 365 |
| HS18-2303_PTCFV | 834 | 1826 | 0.804 | 1124 | 2427 | 721 | 1383 | 0.770 | 1010 | 1921 |
| HS19-4837_PTCFV | 657 | 1901 | 0.786 | 980 | 2615 | 596 | 1767 | 0.811 | 853 | 2334 |
| SS12-3681_PTCFV | 1876 | 3126 | 0.981 | 1960 | 3161 | 1280 | 1997 | 0.975 | 1362 | 2019 |
| SS19-930_PTCFV | 814 | 3731 | 0.790 | 1113 | 5414 | 701 | 3608 | 0.846 | 892 | 4728 |
| SS20-10836_PTCFV | 1167 | 1269 | 0.792 | 1462 | 1639 | 1069 | 1240 | 0.825 | 1298 | 1537 |
| SS20-12603_PTCFV | 1069 | 3848 | 0.770 | 1764 | 5447 | 965 | 3268 | 0.767 | 1567 | 4648 |
| SS20-17729_PTCFV | 656 | 3407 | 0.863 | 756 | 4388 | 551 | 3017 | 0.878 | 643 | 3763 |
| SS20-2973_PTCFV | 1158 | 1891 | 0.595 | 1914 | 3713 | 329 | 662 | 0.877 | 393 | 783 |
| SS20-3963_PTCFV | 2027 | 3154 | 0.895 | 2230 | 3693 | 1922 | 3048 | 0.910 | 2102 | 3472 |
| SS20-9247_PTCFV | 2601 | 3991 | 0.795 | 3228 | 5420 | 2416 | 3525 | 0.800 | 2928 | 4780 |

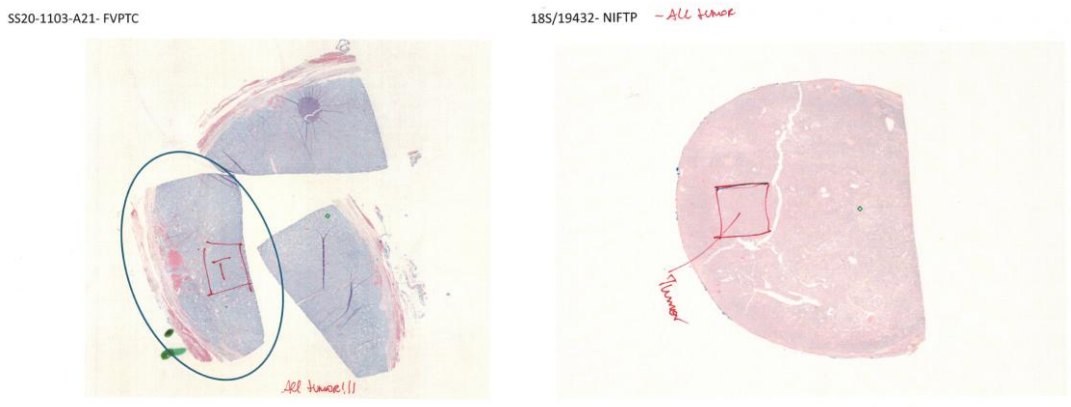
**Supplementary Figure 1:** K-means classification results prior to the "mixed-characteristics" filtering.




**Supplementary Table 2:** K-means classification results prior to the "mixed-characteristics" filtering.

| PTC | | FVPTC | | NIFTP | |
|---|---|---|---|---|---|
| F1 | $\sigma(F1)$ | F1 | $\sigma(F1)$ | F1 | $\sigma(F1)$ |
| 0.823 | 0.057 | 0.807 | 0.094 | 0.769 | 0.066 |

**Supplementary Figure 2:** Examples of cases that contain only tumor regions according to pathologists labeling, making them unsuitable for the semi-automated classification method using K-means clustering (section 2.4).




As explained in the text, this method is envisioned to assist the pathologists as they observe the images and requires them to label on the monitor regions suspected to be normal and tumor, for the computation of the spectral $I_1$ and $I_2$ features. In such cases, where the entire tissue sample expresses tumor features, the characteristics are more profound and easier to detect.

Nevertheless, the automated RFC classification method is still valid, and as we have shown it is used to gain the classification per nucleus (see results section 3).

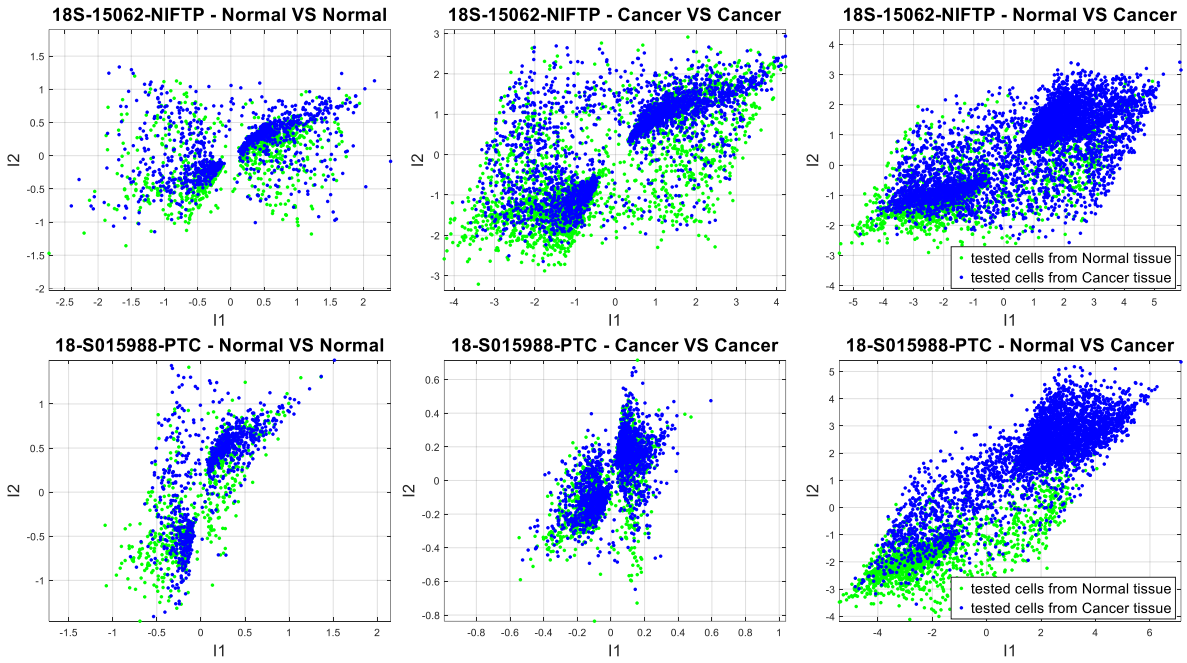
**Supplementary table 3:** K-means clustering results of normal-vs-normal and cancer-vs-cancer, displaying the true-negative and true-positive rates (percentages of nuclei from each group that were correctly classified).

| | Normal-vs-Normal | | | Cancer-vs-Cancer | | |
|---|---|---|---|---|---|---|
| **Case** | **TN-rate** | **TP-rate** | **F1** | **TN-rate** | **TP-rate** | **F1** |
| 18S_15062_NIFTP | 48.285 | 57.236 | 0.544 | 63.375 | 61.995 | 0.619 |
| 18S_44928_NIFTP | 51.103 | 62.799 | 0.604 | 55.564 | 52.885 | 0.574 |
| HS19_12314_NIFTP | 53.834 | 51.668 | 0.493 | 58.692 | 61.270 | 0.629 |
| HS19_18274_NIFTP | 55.758 | 59.639 | 0.575 | 61.914 | 45.320 | 0.497 |
| HS20_13920_NIFTP | 52.161 | 59.329 | 0.597 | 62.237 | 60.198 | 0.617 |
| HS20_21416_NIFTP | 55.714 | 63.329 | 0.587 | 54.621 | 55.749 | 0.566 |
| SS18_1993_NIFTP | 50.993 | 54.991 | 0.524 | 65.541 | 63.657 | 0.644 |
| SS18_199_NIFTP | 45.828 | 60.650 | 0.539 | 45.571 | 64.485 | 0.577 |
| SS18_5274_NIFTP | 54.442 | 56.836 | 0.536 | 66.352 | 55.604 | 0.595 |
| SS19_12451_NIFTP | 54.286 | 69.886 | 0.584 | 54.093 | 49.496 | 0.514 |
| SS19_2344_NIFTP | 56.842 | 59.958 | 0.626 | 58.152 | 55.716 | 0.600 |
| SS19_4159_NIFTP | 51.68 | 53.809 | 0.535 | 51.876 | 54.861 | 0.542 |
| SS19_5490_NIFTP | 58.692 | 51.642 | 0.561 | 54.527 | 60.809 | 0.606 |
| SS19_8215_NIFTP | 59.023 | 57.505 | 0.541 | 56.751 | 51.542 | 0.532 |
| 18_S009289_PTC | 61.150 | 56.826 | 0.563 | 78.852 | 25.756 | 0.351 |
| 18_S013537_PTC | 59.904 | 60.023 | 0.637 | 66.835 | 64.621 | 0.641 |
| 18_S013539_PTC | 58.654 | 52.164 | 0.554 | 63.583 | 49.860 | 0.516 |
| 18_S015988_PTC | 55.140 | 65.960 | 0.642 | 51.251 | 55.156 | 0.558 |
| 18_S020329_PTC | 55.750 | 52.613 | 0.521 | 56.825 | 43.354 | 0.498 |
| SS18_4375_PTC | 47.255 | 58.724 | 0.538 | 75.430 | 72.347 | 0.732 |
| SS18_4748_PTC | 58.96 | 61.806 | 0.586 | 51.498 | 56.166 | 0.528 |
| SS18_5367_PTC | 64.185 | 65.955 | 0.617 | 60.920 | 45.547 | 0.492 |
| SS19_12483_PTC | 50.026 | 54.4 | 0.542 | 51.472 | 54.741 | 0.524 |
| SS19_438_PTC | 54.162 | 54.678 | 0.548 | 62.375 | 57.193 | 0.574 |
| SS19_785_PTC | 69.43 | 72.917 | 0.716 | 56.980 | 51.055 | 0.514 |
| HS18-2303_PTCFV | 59.111 | 54.566 | 0.505 | 62.185 | 62.906 | 0.653 |
| HS19-4837_PTCFV | 63.322 | 60.448 | 0.567 | 51.943 | 59.749 | 0.569 |
| SS12-3681_PTCFV | 59.853 | 59.248 | 0.602 | 64.044 | 64.298 | 0.672 |
| SS19-930_PTCFV | 62.651 | 56.017 | 0.629 | 59.925 | 59.551 | 0.623 |
| SS20-10836_PTCFV | 64.628 | 54.507 | 0.568 | 47.081 | 58.633 | 0.559 |
| SS20-12603_PTCFV | 61.945 | 56.498 | 0.593 | 57.874 | 50.239 | 0.539 |
| SS20-17729_PTCFV | 56.589 | 55.556 | 0.553 | 57.137 | 47.481 | 0.489 |
| SS20-2973_PTCFV | 59.568 | 47.346 | 0.501 | 46.587 | 63.948 | 0.535 |
| SS20-3963_PTCFV | 51.635 | 60.393 | 0.567 | 46.64 | 60.58 | 0.571 |
| SS20-9247_PTCFV | 61.013 | 56.275 | 0.564 | 50.297 | 66.237 | 0.597 |

For each case, the normal nuclei were divided into two groups, approximately equal in size, and were classified using the semi-automated K-means clustering algorithm. The same procedure was applied over the cancer nuclei for each case.

Two identical groups would not be separable and would lead to logical error in our formulas. From these control experiments we can gain understanding of the common levels of variability within each case, which seem to extend up to 0.6 in F1 score with few outliers that may be caused by specific phenomena in these cases, possibly regarding the sample preparation. Another interesting result of this control experiment is the relatively small range of data points in the log-log plots. For the normal-vs-cancer plots, the range of $I_1$ values can be from -4 to 4 (see Fig. 5), while in this control experiment, attempting to compare two normal or two tumor regions, most data points range from -2 to 2, and sometimes much closer to the origin. Furthermore, in many cases it is visible that the masses of data points are located at the 2nd and 4th quartiles, unlike the classification of normal-vs-cancer where they mostly exist in the 1st and 3rd. This is another evidence to the similarity between the control groups, as expected in case of two regions of the same type (normal or tumor).

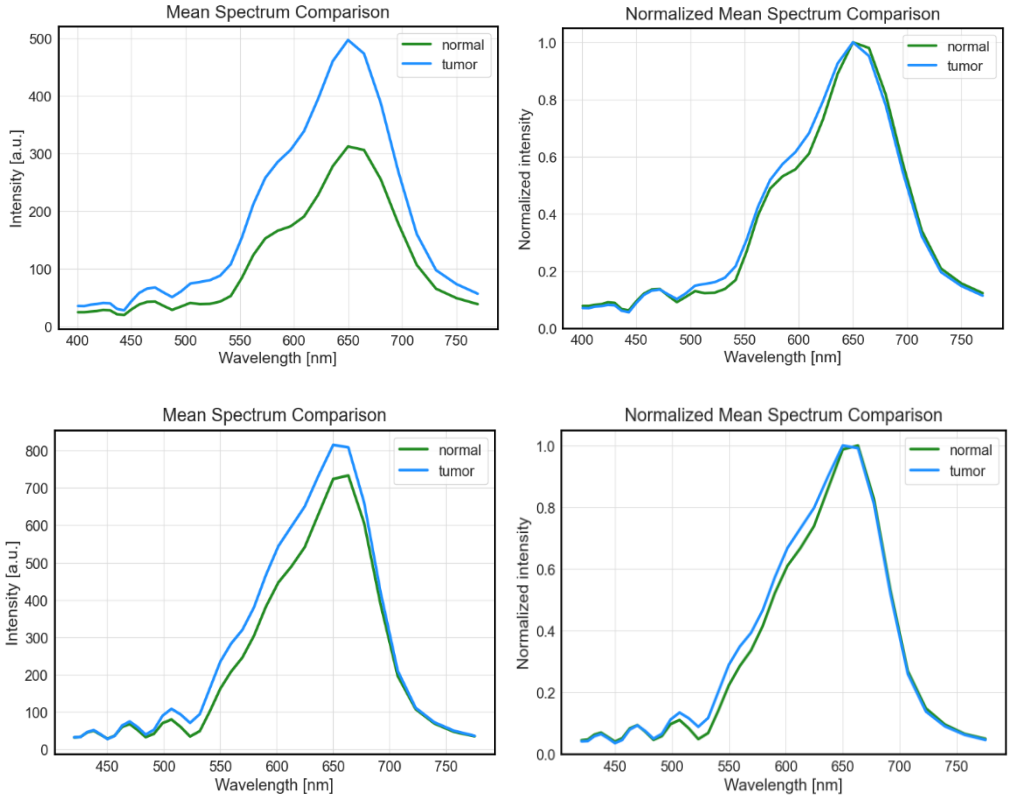
The $I_2$ parameter, that highlights the spectral shape by normalizing the values used in $I_1$, shows a significant spread of the data points when comparing the normal and cancer nuclei. Below, in Supplementary Figure 3, we the results of our K-means method for two cases as an example, the upper row for one case and the lower row for the other. On the right plots are the classification results between the normal and cancer groups, on the middle plots are the classification results for cancer-vs-cancer and on the left plot, normal-vs-normal (see numerical results in supplementary table 3). Clearly, the values on the $I_2$ axis (vertical) are spread much more when comparing the normal and cancer cells rather within each group. This shows that the minor difference seen when observing the spectral shape (Figure 5 in the main text), even if for no more than a few nanometers-shift or intensity difference per wavelength, is captured significantly enough to be utilized. For further observation of the change in spectral shape, Supplementary Figure 4 displays the mean spectrum from Figure 5 compared with the normalized spectrum (first row) and another example to support (Second row), with the expected separations around the absorption wavelengths of hematoxylin and eosin, at 600nm-630nm and 500-550nm correspondingly.




**Supplementary figure 3:** K-means classification results of two example cases, with normal-vs-normal (left column), cancer-vs-cancer (middle column) and normal-vs-cancer (right column).

Another experiment we conducted was to compare the data from normal or cancer regions from different cases, to understand the extent of the variability between different patients. We tested all possible pairs of cases using our semi-automated K-means algorithm and reached average F1 scores of 0.928 for tumor regions and 0.926 for normal regions. This indicates the great variability between data from different patients. However, we suggest that this variability is mainly due to differences in the sample preparation process or in the imaging process that are not intrinsic to the tissue, which is why the automated RFC classification was still able to train and achieve high performance.

**Supplementary figure 4:** comparison of mean spectrum for cancer and tumor nuclei from two different cases, displaying the expected difference in the spectral shape in the absorption wavelengths for hematoxylin and eosin. First row: the example from figure 5 in the main text. Second row: another case displaying the phenomenon.



## 9.3 Sub-Region Classification

**Supplementary table 4:** Region-wise RFC classification results per case in the test dataset, depending on the number of sub-regions.

| Case | Label | Number of images | Correct predictions rate | | |
|---|---|---|---|---|---|
| | | | 1 region per image | 4 regions per image | 6 regions per image |
| SS19_12483_PTC | Cancer | 6 | 0.833 | 0.917 | 0.889 |
| | Normal | 10 | 1 | 1 | 1 |
| SS19_438_PTC | Cancer | 6 | 1 | 1 | 1 |
| | Normal | 7 | 1 | 0.929 | 0.929 |
| SS19_785_PTC | Cancer | 2 | 1 | 0.500 | 0.500 |
| | Normal | 2 | 1 | 0.875 | 0.750 |
| SS20-12603_PTCFV | Cancer | 7 | 0.571 | 0.607 | 0.619 |
| | Normal | 6 | 0.833 | 0.875 | 0.806 |
| SS20-17729_PTCFV | Cancer | 6 | 1 | 0.958 | 0.944 |
| | Normal | 6 | 1 | 0.917 | 0.889 |
| SS19_12451_NIFTP | Cancer | 6 | 1 | 1 | 1 |
| | Normal | 2 | 1 | 1 | 1 |
| SS19_2344_NIFTP | Cancer | 5 | 0.800 | 0.850 | 0.867 |
| | Normal | 6 | 0.875 | 0.875 | 0.833 |
| SS19_4159_NIFTP | Cancer | 7 | 0.714 | 0.786 | 0.690 |
| | Normal | 6 | 0.500 | 0.458 | 0.444 |
| SS19_8215_NIFTP | Cancer | 6 | 1 | 0.833 | 0.806 |
| | Normal | 5 | 1 | 1 | 1 |
| SS19_5490_NIFTP | Cancer | 5 | 1 | 1 | 1 |
| | Normal | 7 | 0.714 | 0.750 | 0.786 |
| SS20-17347_PTCFV | Cancer | 1 | 0 | 0 | 0 |

We divided the tumor and normal regions of each case into smaller sub-regions and tested the effect of the number of sub-regions on the classification results. The entire region is assigned to one of the classes according to the majority class of the nuclei in that region, using the RFC results. As can be seen in the table, the single region reached the highest average result. It is important to remember that a single region is still a limited area on the tissue sample, determined by the size of the acquired image. Further testing can check the effect of even larger images on the results.

The nucleus-wise classification provides verbose information for pathologists; however, it is often excessively detailed for diagnostic purposes. The higher accuracy and confidence levels, together with the ability to adjust the threshold and sub-region size depending on the task at hand, makes sub-region classification of thyroid cancer a useful tool.