European Journal of Radiology 99 (2018) 1–8

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Review

Texture analysis and machine learning to characterize suspected thyroid nodules and differentiated thyroid cancer: Where do we stand?

Martina Sollini[a], Luca Cozzi[a,b], Arturo Chiti[a,c], Margarita Kirienko[a,*]

[a] Department of Biomedical Sciences, Humanitas University, via Rita Levi Montalcini, 20090 Pieve Emanuele (Milan), Italy
[b] Radiotherapy and Radiosurgery, Humanitas Clinical and Research Center, via Manzoni 56, 20089 Rozzano (Milan), Italy
[c] Nuclear Medicine, Humanitas Clinical and Research Center, via Manzoni 56, 20089 Rozzano (Milan), Italy

ARTICLE INFO

Keywords:
Texture analysis
Machine learning
Radiomics
Thyroid nodule
Differentiated thyroid cancer

ABSTRACT

In thyroid imaging, "texture" refers to the echographic appearence of the parenchyma or a nodule. However, definition of the image characteristics is operator dependent and influenced by the operator's experience. In a more objective texture analysis, a variety of mathematical methods are used to describe image inhomogeneity, allowing assessment of an image by means of quantitative parameters. Moreover, this approach may be used to develop an efficient computer-aided diagnosis (CAD) system to yield a second opinion when differentiating malignant and benign thyroid lesions. The aim of this review is to summarize the available literature data on texture analysis, with and without CAD, in patients with suspected thyroid nodules or differentiated thyroid cancer, and to assess the current state of the approach.

1. Introduction

In the United States, 4–7% of the adult population have a clinically identifiable thyroid nodule [1–3]. Although only 5% of palpable nodules are malignant, it is important to identify and treat them properly [1–4]. Imaging techniques such as scintigraphy, ultrasonography (US), computed tomography (CT), and magnetic resonance (MRI) are currently used to evaluate thyroid nodules but they cannot reliably differentiate benign from malignant lesions.

Texture analysis, referring to the characterization of a surface or volume by means of a number of numerical features, may be used in different applied arenas of pattern recognition, including cosmology, art, and medical imaging. Traditionally, in thyroid imaging "texture" refers to the echographic appearence of a nodule, which may be defined as homogeneous or heterogeneous [5], solid, cystic, or mixed [6,7], and hypoechoic, isoechoic, mixed, or hyperechoic [8–12]. However, the definition of these characteristics is operator dependent and influenced by the operator's experience. During recent years, efforts have been made to make evaluation of the echo structure of thyroid nodules more objective (eliminating potential observer bias) in order to standardize reporting, better define the risk of malignancy, and avoid unnecessary biopsies (e.g., TIRADS classification) [13–17]. However, some of the models proposed are difficult to apply in clinical practice and still rely on the observer's assessment of echographic features, even when employing complex methods [18]. Texture analysis allows quantitative measurement of thyroid echogenicity (e.g., the median gray ratio) and echotexture (e.g., coefficient of variation of gray-level histogram), using adimensional features and avoiding the need for fixed US operating conditions (such as depth and gain) [18]. More recently, texture analysis has also been applied to other imaging modalities, such as elastosonography, positron emission tomography/computed tomography (PET/CT), and MRI, to evaluate the risk of malignancy of thyroid nodules or differentiated thyroid cancer (DTC).

Texture features have been used to develop computer-aided diagnosis (CAD) systems. These applications are expected to avoid possible diagnostic errors and to allow shorter and more accurate medical data examination. Diagnosis of thyroid disease can be formulated as a classification problem, so CAD systems are suitable to automatically classify the patient's condition. Machine learning techniques are increasingly being introduced to construct the CAD systems owing to their strong capability to extract complex relationships from the biomedical data [19]. The CAD systems described in the literature use different textural features and machine learning algorithms for classification. Moreover, various parameter optimization techniques are used to achieve better

Abbreviations: ADC, apparent diffusion coefficient; CAD, computer-aided diagnosis; CT, computed tomography; DTC, differentiated thyroid cancer; DWI, diffusion-weighted imaging; [^{18}F]FDG, 2-deoxy-2-(^{18}F)fluoro-D-glucose; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run-length matrix; HF, : heterogeneity factor; MRI, magnetic resonance imaging; PET, positron emission tomography; PTC, papillary thyroid carcinoma; SVM, support vector machine; SWE, shear wave elastography; US, ultrasonography

* Corresponding author. at: Margarita Kirienko, MD, Humanitas University, via Rita Levi Montalcini, 20090 Pieve Emanuele (Milan), Italy.
E-mail addresses: martina.sollini@cancercenter.humanitas.it (M. Sollini), luca.cozzi@cancercenter.humanitas.it (L. Cozzi), arturo.chiti@hunimed.eu (A. Chiti), margarita.kirienko@st.hunimed.eu (M. Kirienko).

https://doi.org/10.1016/j.ejrad.2017.12.004
Received 20 July 2017; Received in revised form 21 November 2017; Accepted 6 December 2017

classification accuracy [20].

This renewed interest in the combination of quantification and visual assessment, able to provide a comprehensive quantification of imaging data sets, may be of great value in thyroid nodule characterization. In fact, the clinical importance of thyroid nodules rests with the need to exclude thyroid cancer, which occurs in 7–15% [4]. The yearly incidence of thyroid cancer nearly tripled from 1975 to 2009 and the percentage of new thyroid cancers ≤1 cm increased from 25% to 39% in 20 years [21]. This tumor shift is due to clinically occult cancers detected incidentally (i.e., as a result of increasing use of neck ultrasonography and other imaging modalities) [22,23]. One study predicts that by 2019 papillary thyroid cancer will become the third most common cancer in women, associated with a cost of $19–21 billion in the United States [24]. Optimization of long-term health outcomes and education about the potential prognosis for individuals with thyroid neoplasms is critically important [4]. Therefore, texture analysis, and machine learning approaches applied to medical imaging may be of value in better characterizing thyroid nodules and in identifying those nodules or tumors with aggressive behavior.

The present report provides a qualitative review of the literature, describing the state of the art in texture analysis methods with and without CAD, in patients with suspected thyroid nodules or DTC.

2. Materials and methods

From the PubMed/MEDLINE database a search algorithm based on a combination of the following terms was used: (a) “texture” or “textural” or “radiomics” or “heterogeneity” or “grey-level” or “gray-level” or “histogram” or “wavelet” or “fractal” and (b) “thyroid”. No start date limit was applied, and the search extended to 22 June 2017. To expand the search, references from the retrieved articles were also screened to identify additional studies. Two authors independently searched articles and performed an initial screening of identified titles and abstracts. All studies or subsets in studies investigating the role of radiomics in patients with suspected/definite DTC were considered eligible for inclusion. Our review was aimed at describing the achievements of texture analysis and CAD applications that could be introduced in routine practice and have an impact on patient management in the near future. Thus, we focused our work on studies performed in the clinical setting since we suppose that the approaches that these studies addressed hold the most promise for implementation. The exclusion criteria were: (a) articles not within the field of interest of this review; (b) articles relating to the development of software for texture analysis; (c) articles dealing with texture analysis of tissue specimens (i.e., cytological/histological images); (d) review or meta-analysis articles, editorials or letters, comments, and conference proceedings; (e) articles not in the English language; (f) case reports or small case series (< 5 patients included); and (g) in vitro or animal studies.

Among the 334 studies identified, 34 articles were selected and used for the qualitative synthesis (Fig. 1). The selected studies were grouped into different sets based on the types of image used to extract features.

In the following, several specific textural features will be named. For the mathematical definition of each of these, we refer readers to the original publications or to a more recent review [25], with the caveat that similar names might be applied to slightly different objects due to an element of weak historical consistency in the definitions of the observables.

3. Results

3.1. Texture analysis of US images in patients with thyroid nodules or DTC

Table 1 summarizes the main results of studies that have focused on texture analysis of radiological images. The majority of papers in this group have aimed to test the ability of CAD systems to assist imagers in differentiating benign from malignant thyroid nodules. This illustrates the growing role that computerized systems play to assist imagers to increase the accuracy of suspected nodule assessment.

The first study on US texture analysis in thyroid nodules at risk of malignancy was published by Hirning et al. in 1989 [26]. Their classification system reached an overall accuracy of 90% in differentiating among classes (i.e., carcinoma, adenoma, struma nodosa, cyst, and normal) [26].

Tsantis et al. [27] developed a support vector machine-based (SVM) image analysis system employing 40 textural features, which performed better than the conventional US evaluation for assessment of the risk of malignancy of thyroid nodules. The SVM identified the set of features with the highest classification accuracy (area under the curve, AUC = 0.97) using the smallest number of textural parameters (i.e., mean, sum variance, and run length non-uniformity). The same group [28] also tested a set of 20 features coupled with two powerful pattern recognition algorithms in order to quantify their power of differentiation, which proved similarly informative.

Savelonas et al. [29] proposed a computer-based approach which provided statistically significant information for the discrimination of thyroid nodules in 171 US images in terms of malignancy risk, contributing to objectification of the diagnostic process [29].

The same group investigated a novel computational approach to thyroid tissue characterization in US images, based on the directionality patterns. Encoding of the directional patterns was realized by means of Radon transform features. The proposed approach was capable of discriminating normal and nodular thyroid tissues (overall accuracy 90.9%), and nodular tissue with high or low malignancy risk (overall accuracy 89.4%) [30].

Iakovidis et al. [31] proposed a novel approach for thyroid US pattern representation that encodes texture and echogenicity features (fuzzy local binary patterns and fuzzy intensity histograms, respectively) via a noise-resistant representation. Overall, the encoded thyroid ultrasound patterns were discriminated by support vector classifiers. The proposed fusion scheme outperformed previous thyroid ultrasound pattern representation methods.

Chen et al. [32,33] analyzed the US images of 61 patients with suspected thyroid nodules with a uniform echo pattern by means of a hierarchical SVM classification system, based on textural features able to differentiate follicle- and fibrosis-based neoplasms, and their subtypes (i.e., follicles with few cells and follicles with dominant cells). The hierarchical SVM classification system achieved a high diagnostic accuracy (96–100%).

Kale et al. [34], and Song et al. [35] evaluated the use of texture-based gray-level co-occurrence matrix (GLCM) features extracted from thyroid US patterns in building prediction models applying different classifiers to determine the nature of thyroid nodules, with positive results.

Grani et al. [18] evaluated 476 patients to develop an objective and reproducible method to stratify the degree of echogenicity (based on histogram analysis) and heterogeneity (coefficient of variation of gray-level histogram) of thyroid nodules in order to predict their risk of malignancy. No significant difference was recorded (using numerical estimates of echogenicity) between iso-hypoechogenic nodules and hypoechogenic, iso-hypoechogenic and isoechogenic, and iso-hypoechogenic and heterogeneous nodules. A real measurable difference was observed between hypoechogenic and isoechogenic nodules. The nodule/parenchyma gray ratio, as an estimation of hypoechogenicity, was able to objectively quantify the degree of hypoechogenicity and estimate the risk of malignancy, with substantial interobserver agreement and diagnostic accuracy. The heterogeneity index of the nodule was increased in malignant lesions.

Ardakani et al. [36] evaluated the ability of a CAD system with texture analysis to improve radiologists' accuracy in identifying thyroid nodules as malignant or benign in 70 patients. Among 270 texture parameters, the ten resulting in the best performance were kurtosis$_{Hist}$, sum entropy$_{GLCM}$, sum of squares$_{GLCM}$, energy$_{W}$ (five components in