**Paper Title**: CLASSIFICATION OF THYROID NODULES IN ULTRASOUND IMAGES USING DEEP MODEL BASED TRANSFER LEARNING AND HYBRID FEATURES

Authors: Tianjiao Liu1 , Shuaining Xie1 , Jing Yu2 , Lijuan Niu3 , Weidong Sun

Venue: New Orleans, LA, USA

|  |  |
| --- | --- |
| File: |  |

URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7952290

Problem: how to capture appropriate features for discriminating benign and malignant nodules in the ultrasound images

Contribution: proposing a feature extraction method for ultrasound images based on the convolution neural networks (CNNs), and trying to introduce more meaningful semantic features to the classification

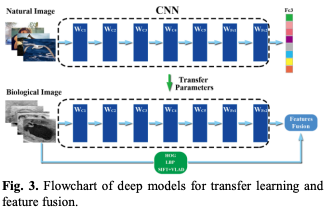
Method/Solution:

* Pre-trained VGG-F Model for Transfer Learning
* Feature fusion and Classification, ada dua macam strategi fusi fitur yang digunakan pada metode yang diajukan pada paper ini:
  + Feature-selected strategy: Standar seleksi fitur berdasarkan peng-urutan nilai perbedaan dari sampel nodul jinak dan ganas. Gambarannya dapat dilihat pada persamaan berikut:



Dengan NMB dan NMM adalah jumlah nodul jinak dan ganas dalam training set, sementara vik adalah fitur dimensi ke-k dari gambar ke-i. 1000 fitur teratas dengan selisih terbesar akan dipilih sebagai fitur final untuk klasifikasi nodul tiroid.

* + Positive-sample-first majority voting strategy. Andaikan terdapat T tipe fitur yang dapat di-ekstraksi dari gambar US Tiroid. Untuk metode ekstraksi fitur k, sebuah “classifier” hk dapat dilatih pada dataset. Nilai yang diperkirakan oleh “classifier” untuk sampel x adalah hk(x). Hasil klasifikasi akhir untuk sampel x dapat diekspresikan dalam notasi sebagai berikut: h(x) = mode (h1(x),...,hT(x)), dengan mode adalah operasi modal. Jika voting kondisi jinak dan ganas bernilai sama, dalam kasus nilai T adalah bilangan genap, maka sampel akan dipertimbangkan sebagai ganas. Begitu voting antara jinak dan ganas bernilai sama terjadi, hasil yang diberikan oleh model ini adalah keganasan, berdasarkan pertimbangan untuk meningkatkan sensitivitas system klasifikasi. Flowchart algoritma metode system yang diajukan dalam paper ini dapat terlihat pada gambar 3:



Data and Evaluation Indexes

Thyroid ultrasound images yang digunakan dalam eksperimen ini diperoleh dari Cancer Hospital of Chinese Academy of Medical Sciences, yang secara klinis telah diverifikasi. Mesin ultrasound yang digunakan dalam eksperimen ini adalah GE Logiq E9, frekuensi detector adalah 10-14MHz. 1037 gambar nodul thyroid, yang terdiri dari 651 gambar jinak dan 386 gambar yang ganas, digunakan dalam eksperimen ini. Tipe dan lokasi nodul di-anotasi oleh dokter. ROI tiap gambar nodul di-ekstraksi melalui demarkasi dokter dan klasifikasi dilakukan menggunakan “classifier” SVM dengan validasi 10-fold cross validation. Indeks evaluasi kuantitatif adalah sbb:

(1) Accuracy= (TP+TN)/(TP+TN +FP+FN);

(2) Sensitivity=TP/(TP+FN);

(3) Specificity =TN/(TN+FP);

(4) Receiver operating characteristic curve (ROC) dan area under curve (AUC).

Dengan TP (true positive) dan TN (true negative) merepresentasikan jumlah hasil klasifikasi positif dan negative yang benar pada sampel. FP (false positive) dan FN (false negative) adalah jumlah klasifikasi yang keliru untuk sampel negative dan positif. Dalam klasifikasi sampel nodul tiroid, sampel positif adalah nodul dengan keganaan, dan sebaliknya (sampel negative adalah nodul jinak).Sensitivitas dan spesifisitas masing-masing mendefinisikan kemungkinan memprediksi suatu nodul memiliki keganasan dan memprediksi suatu nodul jinak.

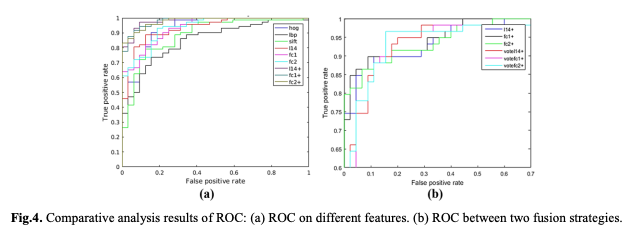
Main Results:

Diperoleh bahwa skor specificity lebih tinggi dari sensitivity dalam tiap fitur yang dipertimbangkan. Hal ini mungkin berasal dari dua alasan:Pertama, jumlah sampel positif pada dataset hampir dua kali lipat dari sampel negatif. Kedua, keragaman tumor ganas membuat perbedaan satu dengan yang lain lebih banyak daripada pada nodul jinak. Selain itu, nodul jinak memiliki karakteristik yang lebih unik, sehingga membuatnya lebih mudah dikenali.

Fitur hybrid yang dihasilkan oleh CNN dapat merepresentasikan karakter inheren dari gambar tanpa memandang tipe gambar, sehingga men-transfer fitur CNN untuk domain gambar ultrasound menjadi fisibel. Selain itu, fitur yang dihasilkan dari layer CNN yang berbeda berperilaku berbeda pula.

AUC bertambah ketika layer semakin tinggi. Pada kondisi koleksi data yang tidak seimbang, AUC dapat secara persuasif memperkuat kinerja pengklasifikasi. Nilai AUC yang relative tinggi mengindikasikan bahwa fc2 adalah grup fitur klasifikasi yang superior. Penjelasan yang dapat diterima dari hal ini adalah layer yang lebih tinggi menangkap fitur yang generic di antara gambar yang berbeda, sementara layer yang lebih rendah hanya menghitung fitur level rendah yang tidak dapat merepresentasikan semantic tingkat tinggi.

Menggabungkan fitur yang berbeda selalu dapat memberikan hasil yang memuaskan. Strategi penggabungan ini memberikan hasil sensitivitas yang secara umum dapat lebih tinggi 11%, dibandingkan dengan metode tanpa penggabungan. Strategi pemilihan fitur termasuk mengkombinasikan dan men-screening fitur, akan menghapus fitur redundan dan fitur yang tidak diinginkan. Berdasarkan perlakuan ini, sekelompok fitur dengan perbedaan maksimal dapat diperoleh. Sementara, untuk mekanisme voting, sensitivitas dapat ditingkatkan dengan berfokus pada pengenalan sampel positif, dan mengadopsi strategi penambahan sampel positif. Perbedaan antara kedua strategi dapat dilihat pada gambar 4(b), strategi pemilihan fitur bertujuan memperbaiki tingkat TP ketika tingkat FP tetap rendah, sementara strategi lain bertujuan memastikan akurasi klasifikasi dari sampel positif.



Limitation: Capaian akurasi pada nodul yang mengandung keganasan masih perlu perbaikan.

Author first name:

Author surname:

Paper Title: Ultrasonographic Characteristics of Thyroid Nodules Prediction of Malignancy

Authors: Eisuke Koike, MD; Shiro Noguchi, MD, PhD; Hiroyuki Yamashita, MD, PhD; Tsukasa Murakami, MD, PhD; Akira Ohshima, MD, PhD; Hitoshi Kawamoto, PhD; Hiroto Yamashita, MD, PhD

URL: https://jamanetwork.com/journals/jamasurgery/fullarticle/391154

Problem: Ultrasonografi real-time (US) resolusi tinggi dapat mendeteksi karakteristik nodul tiroid, tetapi belum dapat menjelaskan dengan baik perbedaan antara nodul ganas dan nodul jinak

Contribution: Karakteristik US untuk memprediksi keganasan baik pada neoplasma folikular maupun nonfolikular melalui analisis regresi logistik ganda dengan menggunakan 5 fitur US.

Method/Solution: Sebuah studi retrospektif dari 329 nodul thyroid (>=5 mm) pada 309 pasien untuk membandingkan karakteristik Ultrasound (US) dan hasil patologi-nya.

Main Results:Sensitivitas diagnosis US pre-operatif adalah 86.5% untuk nonfollicular neoplasms dan 18.2% untuk follicular neoplasms. Spesifisitas adalah 92.3% dan 88.7%, masing-masing untuk nonfollicular neoplasms dan follicular neoplasms. Berdasarkan analisis regresi logistic berganda, margin, shape, echo structure, echo- genisitas, dan kalsifikasi adalah indikasi yang handal untuk keganasan di nonfollicular neoplasms. Berdasarkan kurva “receiver operating characteristic” (RoC) yang dihasilkan dari analisis regresi logistic berganda, titik terbaik untuk tidak mengabaikan keganasan adalah ketika sensitivitas 94% dan spesifisitas 87%. Peluang keganasan pada titik ini lebih besar dari 0.2. Sementara untuk Non-Follikular Neoplasma, diagnosis US tidak reliabel, bahkan saat analisis regresi logistic berganda diterapkan.

Limitation: Formula yang diajukan dalam paper ini dapat diterapkan pada nonfollicular thyroid neoplasma dengan akurasi tinggi, namun tidak dapat diterapkan pada follicular thyroid neoplasma.

Paper Title: Deep learning in Ultrasound Imaging

Authors: RUUD J.G. VAN SLOUN1, REGEV COHEN2 AND YONINA C. ELDAR3

URL: https://arxiv.org/pdf/1907.02994.pdf

Problem: **possible impact dari metodologi deep learning pada banyak aspek dari ultrasound imaging**

Contribution: **deep learning strategies in ultra- sound systems, from the front-end to advanced applications**

Method/Solution: Pada paper ini penulis menampilkan rancangan beberapa “building blocks” terpisah, dengan agen artifisial terlatih dan pemroses sinyal saraf yang khusus untuk aplikasi tertentu. Sebagian metode beroperasi pada gambar atau data IQ sementara yang lain memproses kanal data secara langsung. Proposisi ini memungkinkan rantai pencitraan yang khusus ditujukan untuk aplikasi dan bersifat adaptif sepenuhnya.

Main Results:

Harapan yang dapat diberikan deep learning untuk pencitraan ultrasound cukup signifikan; ini dapat mendorong pergeseran paradigma dalam perancangan system ultrasound systems, di mana probes nirkabel cerdas difasilitasi oleh sub-Nyquist dan tepi-saraf computing dihubungkan pada cloud, dan dengan algoritma serta mode pencitraan yang dikendalikan oleh AI untuk aplikasi tertentu.

Diperkaya dengan deep- learning, pencitraan ultrasound generasi selanjutnya dapat menjadi modalitas yang lebih kuat dengan perangkat yang secara kontinyu belajar untuk memberikan gambar dan pandangan klinis lebih baik, sehingga membawa diagnostic yang lebih baik dan mudah untuk diakse serta pencitraan cerdas yang memiliki portabilitas tinggi.

Limitation:

Author first name:

Author surname:

|  |  |
| --- | --- |
| Paper Title: Comparison Study of Radiomics and Deep Learning-Based Methods for Thyroid Nodules Classification Using Ultrasound Images  File: |  |
| Authors : YONGFENG WANG1, WENWEN YUE2, XIAOLONG LI2, SHUYU LIU3, LEHANG GUO2,  HUIXIONG XU2, HEYE ZHANG 4, (Member, IEEE), AND GUANG YANG 5, (Member, IEEE) |  |

URL:

Problem: Belum ada komparasi antara pendekatan berbasis radiomics based dan deep learning dalam melakukan klasifikasi pada nodul tiroid dari gambar ultrasound

Contribution:

Method/Solution: Untuk sisi radiomik, penulis mengembangkan metode berbasis radiomic, yang terdiri dari ekstraksi 302 fitur statistic dimensional ber-troughput tinggi dari gambar yang telah di-preproses. Kemudian pengurangan dimensi dilakukan dengan informasi mutual dan analisis diskriminan linear untuk menghasilkan klasifikasi akhir. Di sisi lain, metode berbasis deep learning juga dikembangkan dan diuji oleh model VGG16 yang telah dilatih sebelumnya. Gambar ultrasound yang terdiri dari 3120 citra (1841 nodul jinak dan 1393 nodul ganas) dari 1040 kasus dikumpulkan secara retrospektif. Dataset dibagi menjadi 80% untuk training dan 20% untuk pengujian

Main Results: Metode berbasis deep learning dapat memiliki performansi yang lebih baik daripada metode yang menggunakan radiomic. Akurasi yang dihasilkan dari data testing untuk metode berbasis radiomic dan deep learning masing-masing 66.81% dan 74.69%.

Limitation: Metode berbasis deep learning yang digunakan dalam paper ini hanya model CNN umum (VGG16) dan dilatih dengan dataset yang kecil. Sementara berdasarkan studi sebelumnya, pengaruh ukuran dataset pada transfer learning menggunakan CNN dapat memperbaiki unjuk kerja model deep learning.

Author first name:

Paper Title : Management of Thyroid Nodules Seen on US Images: Deep Learning May Match Performance of Radiologists

|  |  |
| --- | --- |
| File: |  |

URL: https://pubs.rsna.org/doi/pdf/10.1148/radiol.2019181343

Problem: Manajemen nodul tiroid mungkin tidak konsisten antara pengamat yang berbeda dan memakan waktu untuk ahli radiologi. Sistem kecerdasan buatan yang menggunakan deep learning dapat meningkatkan alur kerja radiologi untuk pengelolaan nodul tiroid

Contribution: algoritma deep learning yang menggunakan gambar ultrasound (US) untuk memutuskan apakah suatu nodul tiroid perlu diberi tindakan biopsy dan untuk membandingkan unjuk kerja algoritma terhadap yang dilakukan radiolog yang mematuhi prosedur standar American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS).

Method/Solution: Dalam analisis retrospektif, studi terhadap pasien yang dirujuk untuk US dengan tindakan lanjutan “fine-needle aspiration” (FNA) atau analisis histologis bedah sebagai standar dievaluasi. Periode penelitian ini dilakukan dari Agustus 2006 sampai May 2010. Sebuah multitask deep convolutional neural network dilatih untuk memberikan rekomendasi biopsy pada nodul tiroid berdasarkan dua gambar US orotgonal sebagai input. Pada fase training, algoritma deep learning mula-mula di-evaluasi dengan 10-fold cross-val-idation. Validasi internal kemudian dilakukan pada sekumpulan 99 nodul terpisah. Sensitivitas and spesifisitas algoritma ini dibandingkan terhadap consensus dari tiga komite pakar ACR TI-RADS dan Sembilan radiolog lain, yang menginterpretasikan gambar US thyroid dalam praktik klinis. Data yang disertakan adalah 1377 data gambar lengkap disertai diagnosis histologi atau kesimpulan sitologi dari 1230 pasien.

Main Results: Untuk 99 nodul uji, algoritma deep learning yang diajukan mencapai 13 dari 15 (87%: 95% confidence interval [CI]: 67%, 100%) sensitivitas, sama dengan consensus ahli (P . .99) dan lebih tinggi dari 5 dari 9 radiolog. Spesifisitas dari algoritma deep learning adalah 44 dari 84 (52%; 95% CI: 42%, 62%), yakni sama dengan consensus ahli (43 dari 84; 51%; 95% CI: 41%, 62%; P = .91) dan lebih tinggi dari 7 dari 9 radiolog lain. Sensitivitas dan spesifisitas rata-rata dari 9 radiolog masing-masing adalah 83% (95% CI: 64%, 98%) dan 48% (95% CI: 37%, 59%).

Limitation: tidak ada set tes skala besar dari institusi eksternal yang tersedia untuk perbandingan dan untuk menilai generalisasi ke populasi pasien dan nodul yang lebih luas

Author first name:

Author surname: