

Nama: Niko Muhamad Fajar

NIM: 212410101032

Dataminng A

Tugas Review metode ensemble learning

Artikel: An Analysis on Ensemble Learning Optimized Medical Image Classification With Deep Convolutional Neural Networks

Penulis: DOMINIK MÜLLER, IÑAKI SOTO-REY, AND FRANK KRAMER

1. Data yang digunakan sebagai input pada artikel tersebut ialah dataset CHMNIST, COVID, ISIC, DRD. Atribut yang digunakan meliputi jenis gambar seperti gambar CT-Scan, Sinar-X, Gambar histologi, dermoscopy, dan ophthalmoscopy. Parameter yang digunakan yaitu Jumlah layer, Ukuran filter, Jumlah filter, Ukuran stride, Dropout, Batch size, Metrik evaluasi. Tipe data tersebut yaitu ordinal seperti gambar medis, yg terdiri dari tekstur, bentuk, intensitas, frekuensi, atau distribusi warna. Kelas pada dataset CHMNIST ada 8, COVID ada 3, ISIC ada 8, DRD ada 5.
2. A. Convolutional Neural Networks (CNN): Metode ini merupakan salah satu teknik utama yang digunakan dalam pembelajaran pada citra medis.  
B. Ensemble Learning: Metode ini melibatkan penggabungan beberapa model pembelajaran untuk meningkatkan kinerja klasifikasi. Metode ini terdiri dari Augmentation, Stacking, Boosting, dan Bagging  
C. Transfer Learning: Metode ini memanfaatkan pengetahuan yang telah dipelajari oleh model yang sudah terlatih pada dataset yang berbeda.  
D. Preprocessing dan Normalisasi: Metode ini melibatkan tahap pra-pemrosesan data seperti normalisasi intensitas piksel, cropping, resizing, atau penghapusan noise.
3. A. Normalisasi intensitas piksel: Normalisasi intensitas piksel dilakukan untuk mengubah rentang intensitas piksel menjadi rentang yang konsisten.  
B. Cropping (Pemangkasan): Pemangkasan dilakukan untuk memotong atau membuang bagian tidak penting dari gambar.  
C. Resizing (Pengubahan ukuran): Pengubahan ukuran dilakukan untuk mengubah dimensi gambar menjadi ukuran yang konsisten.  
D. Penghapusan noise: Penghapusan noise dilakukan untuk mengurangi atau menghilangkan noise atau sinyal yang tidak diinginkan dari gambar.  
D. Augmentasi data: Augmentasi data melibatkan pembuatan variasi data tambahan dengan memanipulasi gambar asli. Teknik augmentasi data seperti rotasi, pemindahan, pergeseran, zoom, flipping horizontal/vertikal, atau penambahan gangguan (perturbations) dapat digunakan untuk meningkatkan keberagaman data pelatihan dan meningkatkan keakuratan model.
4. Output dari penelitian tersebut yaitu bahwa Stacking mencapai kinerja terbesar keuntungan hingga 13% peningkatan skor F1. Augmenting menunjukkan kemampuan peningkatan yang konsisten hingga 4% dan juga berlaku untuk jaringan pipa berbasis model tunggal. Bagging berdasarkan validasi silang didemonstrasikan perolehan kinerja yang signifikan mendekati Stacking, yang menghasilkan peningkatan skor F1 hingga +11%. Selain itu, kami mendemonstrasikan bahwa fungsi penyatuan statistik sederhana sama atau seringkali bahkan lebih baik daripada fungsi penyatuan yang lebih kompleks. Kami menyimpulkan bahwa integrasi teknik pembelajaran ansambel adalah sebuah metode yang ampuh untuk pipa klasifikasi citra medis apa pun untuk meningkatkan ketahanan dan meningkatkan kinerja.
5. Kelebihannya adalah bahwa Stacking mencapai peningkatan kinerja terbesar dalam alur klasifikasi citra medis. Augmenting menunjukkan peningkatan kemampuan yang konsisten pada model yang tidak mengalami overfitting dan memiliki keunggulan untuk dapat diterapkan pada jaringan pipa berbasis model tunggal. Bagging berbasis cross-validation juga menunjukkan peningkatan kinerja yang signifikan yang mendekati Stacking, tetapi bergantung pada pengambilan sampel dengan representasi fitur yang memadai di semua lipatan. Selain itu, penulis menunjukkan bahwa fungsi penyatuan statistik sederhana seperti Mean atau Pemungutan Suara Mayoritas sama baik atau bahkan lebih baik daripada fungsi penyatuan kompleks seperti Support Vector Machines. Secara keseluruhan, penulis menyimpulkan bahwa integrasi teknik pembelajaran ensemble adalah metode yang kuat untuk meningkatkan klasifikasi citra medis dalam pipa MIC.