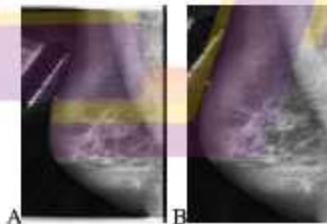


BAB V PENUTUP

5.1 Kesimpulan

Model CNN usulan yang dibangun menggunakan data uji berukuran 200×300 , *learning rate* adam 0.001, jumlah epoch 20, dan menggunakan *average* dan *max pooling* layer mampu mencapai akurasi training hingga 100%, akurasi data validasi 90.46% dan 55% pada data baru. Dengan melakukan beberapa pengujian optimasi, akurasi terhadap data validasi (evaluasi) tertinggi yang dapat dicapai adalah sebesar 91.67% dan akurasi maksimal terhadap data test baru sebesar 60.83%.

Terdapat beberapa hal yang mempengaruhi performa, jumlah layer CNN, ukuran *matrix* kernel, *hyperparameter*, fungsi aktivasi, fungsi loss, jumlah *epoch* (iterasi) dan *batch size*, hingga *optimizer*. Faktor lainnya adalah data dan informasi yang diberikan untuk menunjang performa. Perlu diketahui bahwa kanker merupakan jaringan *abnormal* dan setiap citra mamografi memiliki ukuran massa jaringan yang berbeda. Maka informasi mengenai massa jaringan bisa ditambahkan sebagai fitur tambahan pada penelitian selanjutnya.



Gambar 5.1. Label pada citra mamografi.

Dataset yang digunakan juga memiliki *noise* berupa label pada sebagian *sample* data yang terletak di sudut citra (Gambar 5.1.A). Label ini tentunya mempengaruhi kinerja model dalam melakukan klasifikasi. Sehingga dilakukan penambahan data dengan melakukan augmentasi berupa zoom dengan tujuan mengurangi *noise* (Gambar 5.1.B). Namun akhirnya upaya ini juga belum mampu meningkatkan performa model yang dibangun.

Menggunakan beberapa penyeruaian data, performa arsitektur dan model yang dibangun masih belum mampu melakukan generalisasi informasi dari data yang disediakan. Berdasarkan hasil akurasi data *training*, validasi, dan test, model cenderung menghafal data uji yang diberikan dan belum mampu mengekstrak fitur yang dibutuhkan untuk mengkategorikan data berdasarkan label yang diberikan. Terhadap beberapa optimasi yang telah diberikan, performa model cenderung stabil.

Pada pengujian jumlah epoch dapat disimpulkan bahwa semakin rendah nilai *batch size*, nilai akurasi akan sedikit semakin tinggi. Berbanding terbalik dengan nilai loss yang semakin kecil. Tetapi performa terhadap *batch size* juga ditentukan oleh epoch. Semakin tinggi nilai epoch maka nilai akurasi data training cenderung terus meningkat dengan nilai loss semakin menurun. Tetapi pada data validasi, nilai akurasi akan landai atau cenderung mengalami penurunan pada jumlah epoch tertentu. Hal ini juga mengakibatkan loss yang meningkat pada jumlah epoch tertentu. Parameter *learning rate* yang mampu meningkatkan kualitas model adalah 0.001 baik pada optimizer adam maupun RMSProp.

Pengujian penyesuaian jumlah layer memberikan *output* jumlah layer paling optimal adalah 3 layer dengan *pooling layer max pooling layer* dan activation function relu yang diuji menggunakan 3 skema. Skema pertama, menggunakan optimizer adam, *learning rate* 0.001, dan *epoch* 20, model mampu memberikan akurasi terhadap data test baru adalah sebesar 55%. Nilai *recall*, kemampuan membaca data kanker, dan nilai *precision*, kemampuan membaca data normal, sama yaitu 55%. Performa model terhadap kedua label data seimbang. Berikut dengan nilai F1 sebesar 55%.

Skema berikutnya, menggunakan optimizer adam, *learning rate* 0.001, dan *epoch* 10 model mampu memberikan akurasi terhadap data test baru sebesar 60.83% , nilai akurasi tertinggi diantara model dan skema yang telah diuji, dengan nilai F1 54.37% cenderung bias terhadap label negatif. Skema terakhir menggunakan optimizer RMSprop pada *learning rate* dan *epoch* yang sama, model mampu memberikan akurasi 55% dengan nilai F1 63%, nilai false terhadap kedua label paling rendah. Dengan kemampuan membaca data kanker hingga 76.67%.

5.2 Saran

Berdasarkan dengan hasil penelitian, pembahasan, serta kendala yang dialami oleh penulis selama mengerjakan penelitian, saran yang dapat dipertimbangkan untuk penelitian selanjutnya adalah

1. Penelitian selanjutnya diharapkan memberikan fitur tambahan untuk menambah informasi data citra mamografi, contohnya seperti nilai densitas jaringan payudara, posisi kanker atau masking posisi jaringan abnormal.
2. Menggunakan metode lain atau menggabungkan beberapa metode untuk mencapai performa model dan hasil klasifikasi terbaik.
3. Melakukan analisis terhadap parameter lain seperti *dropout*, *kernel initializer*, ukuran kernel, fungsi aktivasi, atau parameter lain yang disesuaikan dengan data serta metode yang digunakan.
4. Melakukan analisis terhadap objek secara mendalam sehingga mengenali data yang akan diuji.
5. Citra mamografi merupakan citra medis yang mampu membawa informasi penunjang diantaranya berupa ukuran massa jaringan dan mask pada jaringan abnormal. Pada penelitian berikutnya peneliti diharapkan dapat memberi informasi tambahan yang akan menjadi fitur berguna bagi model CNN yang dibangun.