

Age Estimation Untuk *Intelligent Advertising* Pada Poster Digital Menggunakan *Convolutional Neural Network*

^{*1}Galih Abdul Muhyi, ²Yaya Wihardi, ³Erna Piantari

Program Studi Ilmu Komputer, Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam
Universitas Pendidikan Indonesia

Jl. Dr. Setiabudhi No. 229 Bandung 40154 Jawa Barat – Indonesia

¹galih91@student.upi.edu

²yayawihardi@upi.edu, ³erna.piantari@upi.edu

Abstrak— Sebagai bagian dari *intelligent advertising*, *age estimation* digunakan untuk menyesuaikan iklan dari hasil estimasi usia audience. *Age estimation* (AE) dapat dibangun menggunakan deep learning menggunakan ConvNet dengan kendala seperti data training wajah usia tua yang sedikit dan ketidak seimbangan dataset di dalamnya serta membutuhkan jumlah data yang besar. Salah satu solusi dari permasalahan ini adalah melakukan data augmentasi menggunakan model generatif ACGAN untuk melakukan generate gambar sesuai dengan kelas. *Intelligent advertising* pada poster digital hanya disimulasikan pada komputer. Simulasi *intelligent advertising* berfungsi dengan baik terlepas dari terbatasnya iklan dan tidak konsistennya hasil estimasi usia. Hasil dari penggunaan model generatif ACGAN untuk data augmentation berhasil meningkatkan performa hasil pada model AE terlepas dari rendahnya skor IS dan FID serta kualitas gambar yang dihasilkan. Hasil data augmentation lebih terlihat pada model B dengan peningkatan akurasi *cumulative score* sebesar 4,8% dan skor MAE sebesar 1,297.

Kata kunci—*Intelligent Advertising*, *Age Estimation*, ACGAN, *Convolutional Neural Network*.

I. PENDAHULUAN

Poster digital adalah salah satu media periklanan luar rumah modern yang menggunakan layar digital untuk menampilkan berbagai konten [1] seperti informasi, iklan, dan lain sebagainya secara dinamis dan atraktif [2]. Poster digital dapat ditemukan di tempat umum seperti pasar swalayan, stasiun, bandara, dan lain sebagainya [1]. Seiring berkembangnya teknologi, poster digital dapat ditingkatkan dengan menambahkan AI untuk menyesuaikan iklan yang pantas berdasarkan data demografik dari audience seperti usia dan jenis kelamin [2] untuk meningkatkan perhatian dari audience [3]. Pengalaman periklanan luar ruangan tersebut dikenal dengan *intelligent advertising* atau *smart advertising* [4].

Sebagai bagian dari *intelligent advertising*, *age estimation* digunakan untuk menyesuaikan iklan dari hasil estimasi usia audience. *Age estimation* (AE) dapat dibangun menggunakan deep learning sehingga tidak ada campur tangan manusia dalam ekstraksi dan pemilihan fitur [5]. AE

juga digunakan untuk *electronic customer relationship management* (ECRM) dan akses control [5]. Salah satu kendala dalam AE ialah data training wajah usia tua yang sedikit dan ketidak seimbangan dataset di dalamnya. Selain itu, AE yang dibangun menggunakan deep learning memiliki kelemahan pada besarnya jumlah data yang dibutuhkan. Salah satu solusi dari permasalahan ini melakukan data augmentasi menggunakan model generatif ACGAN untuk melakukan generate gambar sesuai dengan kelas yang diinginkan. Data augmentasi menggunakan model generatif sendiri pernah dilakukan oleh [6] menggunakan cycle GAN, [7] dengan DAGAN-nya, dan [8] menggunakan ACGAN. Kelebihan ACGAN yang bisa melakukan generate gambar sesuai label yang diinginkan untuk data augmentation menjadikan motivasi penulis untuk mengajukan solusi untuk permasalahan AE yang dikenal selalu memiliki kekurangan pada data lanjut usia. Hasil dari model AE yang dibangun menggunakan CNN kemudian dilakukan perbandingan akurasi kemudian saat data testing dilakukan pengukuran *cumulative score* dan MAE antara model tanpa data augmentation dan model dengan data augmentation.

II. PENELITIAN TERKAIT

Intelligent Advertising telah diterapkan di berbagai media periklanan luar ruangan seperti billboard atau disimulasikan pada komputer seperti yang dilakukan oleh [9], [10], [11]. Penelitian oleh [9] disimulasikan pada computer menggunakan *skin color detection* selain *face recognition* dengan interaksi berupa perubahan konten saat audience semakin mendekat ke layar. Penelitian oleh [10] menggunakan *face detection* dan *object detection* untuk mendeksi aksesoris yang diuji menggunakan foto untuk menyesuaikan iklan berdasarkan aksesoris yang digunakan. Sedangkan penelitian oleh [11] sudah diimplementasikan pada billboard menggunakan berbagai deteksi seperti *face recognition*, *age estimation*, dan *object detection*.

Penggunaan model ACGAN sebagai data augmentation yang dilakukan oleh [8] berhasil meningkatkan akurasi klasifikasi terhadap dataset dengan data tidak seimbang

dengan 1ma kelas sebesar 5% dan 49,33% apabila menggunakan 3 kelas dengan sampel yang dikurangi.

Age estimation yang diteliti oleh [12], [13], dan [14] menghasilkan nilai MAE dengan nilai satu digit dengan akurasi diatas 50%. Penelitian [12], [13], dan [14] menggunakan dataset populer AE yang tercatat oleh [5]. Penelitian ini menggunakan dataset milik [15] yang bukan didesain untuk AE sehingga diperlukan pengolahan untuk menyesuaikan dengan kebutuhan eksperimen.

III. PEMBANGUNAN SISTEM

Dalam membangun *intelligent advertising* pada poster digital, perlu diperhatikan bahwa sistem hanya disimulasikan di komputer dengan alur pembangunan sebagai berikut.

A. Pengelolaan Dataset

Dataset milik [15] terdiri dari 13322 gambar wajah seperti Gambar 1. Dataset kemudian dilakukan reduksi resolusi gambar dan disimpan tanpa perubahan nama dalam format .png. Dataset yang sudah dilabeli terdiri dari 79 kelas dengan usia dari 2-80 tahun. Dastaset kemudian dilakukan downsampling untuk menyeimbangkan jumlah data yang sudah dilabeli antara usia muda dan usia tua. Dengan demikian, jumlah data menjadi 9258 dengan data setiap kelas tidak lebih dari 200 data. Dataset setelah downsampling diduplikasi untuk dibagi dengan proporsi 80% untuk data training yang dikenal dengan dataaet A untuk training model AE dengan jumlah 7416 data dan 20% untuk data testing model AE dengan jumlah 1842 data.



Gambar 1 Sampel Wajah dari All-Age-Faces Dataset [15]

B. Pembangunan Model Generatif

Model generatif dibangun menggunakan arsitektur ACGAN *vanilla* untuk ImageNet oleh [16] yang dikenal dengan hyperparameter A.. Proses training menggunakan dataset setelah downsampling untuk dengan jumlah epoch 500, 2500, dan 3500 menggunakan *latent space* berukuran 100, 150, dan 200. Sedangkan hyperparameter B dibuat berdasarkan arsitektur *vanilla* dengan perubahan salah satu nilai peta fitur layer transposed convolution pada layer ketiga menjadi berukuran 96. Sebelum memasuki proses training discriminator, resolusi dataset akan direduksi menjadi 128 x 128. Resoulusi gambar yang dihasilkan generator sama dengan resolusi gambar pada discriminator yaitu 128 x 128. Model generatif dengan hyperparameter A dan hyperparameter B akan dievaluasi secara kuantitatif menggunakan IS score dan FID sedangkan secara kualitatif menggunakan kemampuan mata secara manual oleh penulis.

C. Proses Data Augmentation

Proses data augmentasi dilakukan dengan melakukan *generate* gambar dari model terbaik yang sudah dinilai

secara kuantitatif dan kualitatif. Jumlah data yang akan di-generate sebesar 1000 data untuk setiap kelas. Setelah data sudah ditambahkan, dilakukan pembersihan gambar data augmentasi yang berkualitas buruk sesuai dengan jumlah data awal. Dengan demikian dataset B terdiri dari 1000 data untuk setiap kelas dengan total 79000 data.

D. Pembangunan Model Age Estimation

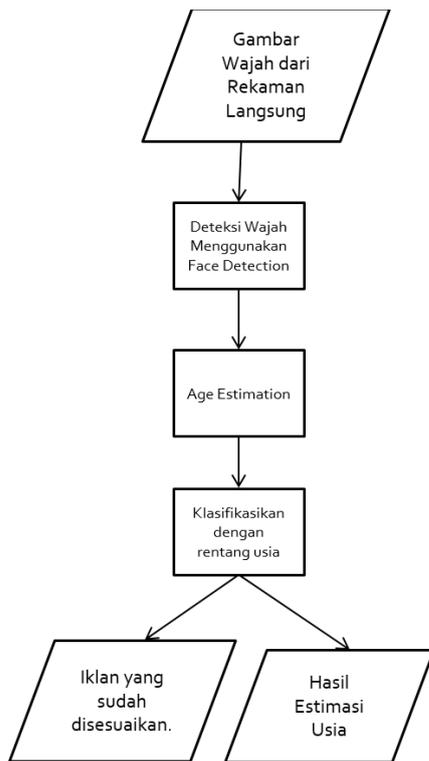
Arsitektur ConvNet pada pembangunan *model age estimation* menggunakan hyperparameter dari discriminator ACGAN dengan penyesuaian sebagai berikut.

1. Layer konvolusi terdiri dari lima layer dengan ukuran peta fitur 32, 64, 128, 256 untuk arsitektur A, Sedangkan arsitektur B menggunakan ukuran peta fitur yang sama seperti arsitektur A dengan penambahan kedalaman layer dengan ukuran peta fitur 512.
2. Seluruh layer menggunakan aktivasi LeakyReLU dengan slope 0,2 dan nilai dropout sebesar 0,25 kecuali untuk arsitektur B pada layer terakhir menggunakan nilai dropout sebesar 0,5.
3. Penambahan max pooling pada akhir layer pertama saja untuk arsitektur A sedangkan arsitektur B menggunakan max pooling pada akhir layer pertama dan kedua.
4. Kedua arsitektur menggunakan optimizer ADAM dengan parameter $\alpha = 0.0002$, $\beta_1 = 0.5$, dan $\beta_2 = 0.999$.
5. Input gambar saat training sebelumnya dilakukan reduksi resolusi menjadi 128 x 128.
6. Kedua arsitektur menggunakan layer linear dengan aktivasi *softmax* dengan 81 peta fitur.

Proses training dilakukan dalam 5000 epoch dengan jumlah satu batch sebesar 1000. Setelah model terbentuk dilakukan evaluasi menggunakan data testing untuk dilakukan perhitungan skor MAE dan *cumulative score* untuk model arsitektur A menggunakan dataset A dan dataset B serta model arsitektur B menggunakan dataset A dan dataset B. Dengan demikian ada empat model yang akan dilakukan pengujian dengan data testing.

E. Pembangunan *Intelligent Advertising*

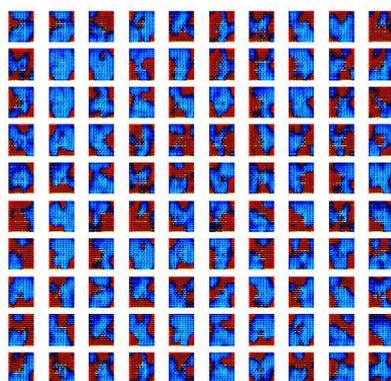
Alur kerja aplikasi dapat dilihat pada Gambar 2. Aplikasi menerima input berupa rekaman wajah kemudian dilakukan face detection untuk mengambil gambar wajah yang terdeteksi. Seluruh gambar wajah terdeteksi akan di-crop kemudian dimasukkan ke dalam model age estimation untuk mendapatkan estimasi usia pada wajah. Data usia akan digunakan untuk menyesuaikan iklan yang akan muncul atau diganti. Output dari aplikasi ini terdiri dari hasil estimasi usia beserta lokasi wajah pada face detection, dan iklan yang sudah disesuaikan.



Gambar 2 Alur Kerja Intelligent Advertising
IV. HASIL DAN PEMBAHASAN

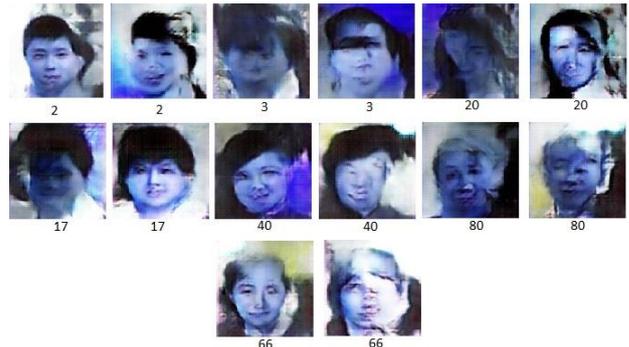
A. Model Generatif

Hasil skor IS dan FID dapat dilihat pada Tabel 1. Ditinjau dari hasil Tabel 1, Skor IS dan FID yang dihasilkan mencapai 1 dan 400 yang menunjukkan rendahnya kualitas sampel gambar yang dihasilkan dari model generatif. Rendahnya kualitas gambar tidak lepas dari terlalu banyaknya jumlah kelas yang digunakan saat training sesuai hasil [16]. Skor terbaik diraih oleh hyperparameter A dengan *latent space* 100 epoch ke-3500 dengan IS dan FID sebesar 1,0029434 dan 471,171. Apabila ditinjau dari kualitas gambar yang dihasilkan, model hyperparameter B dengan *latent space* 150 epoch ke-490 menghasilkan kualitas gambar yang lebih baik dari model yang sama pada epoch ke-500 terlepas dari selisih skor yang dihasilkan seperti Gambar 5. Sampel gambar dari model *hyperparameter A latent space* 100 pada Gambar 3 menunjukkan meskipun hasil skor IS dan FID tinggi tidak menjamin hasil sampel akan sama sehingga diperlukan pengamatan manual.



Gambar 3 Sampel Gambar *hyperparameter A latent space* 100 epoch ke-500

Gambar 4 menampilkan sampel gambar yang di-generate sesuai kelas dari model *hyperparameter B latent space* 150 epoch ke-490 sebagai data augmentation untuk pembangunan dataset B. Sampel gambar dari data augmentation pada kelas usia 2 tahun terlihat seperti gabungan beberapa wajah tanpa memerhatikan kemiripan pada kelas usia 2 tahun pada dataset. Usia 80 tahun tampak mirip seperti penggabungan tidak sempurna dari beberapa wajah dari beberapa kelas usia muda. Dengan mempertimbangkan terbentuknya fitur wajah, dipilihlah model *hyperparameter B latent space* 150 epoch ke-490 sebagai model yang digunakan untuk proses data augmentation.



Gambar 4 Sampel Gambar Hasil Data Augmentation

Tabel 1 Hasil Skor IS dan FID Model Generatif ACGAN

Model Generatif	Epoch	Skor IS ± Deviasi	FID
Hyperparameter A, Latent space 150	500/500	1.0018967 ± 3.096501e-05	472.060
Hyperparameter A, Latent space 100.	3500/3500	1.0029434 ± 5.4418277e-05	471.171
Hyperparameter B, Latent space 150	490/500	1.0017908 ± 3.379433e-05	471.728
Hyperparameter B, Latent space 150	500/500	1.0009428 ± 2.0285835e-05	471.632
Hyperparameter B, Latent space 200.	2500/2500	1.0011818 ± 1.8007193e-05	474.351

B. Model Age Estimation

Grafik dari hasil dari training seluruh model AE dapat dilihat pada Gambar 6 yang menghasilkan akurasi diatas 90%. Penulisan model disingkat menjadi dua huruf dengan singkatan dari arsitektur dan dataset yang digunakan seperti model AA yang berarti model arsitektur A dataset A. Model BA menghasilkan akurasi sebesar 92,7% dan meningkat menggunakan dataset B atau model BB menjadi 99,17%. Akurasi pada model B meningkat sebesar 6,47%. Sedangkan model AA menghasilkan akurasi sebesar 98,26% dan meningkat menggunakan dataset B atau model AB menjadi 99,83%. Akurasi pada model A meningkat sebesar 1,57%. Dilihat dari hasil training model A memiliki hasil lebih baik daripada model B.

Hasil dari uji data testing setiap model dapat dilihat pada Tabel 2 dengan skor MAE beserta CS pada batas 10 tahun dan Tabel 3 yang menampilkan skor CS pada batas error 5, 10, 15, 20, dan 25 tahun. Tabel 2 membuktikan bahwa model B menghasilkan performa lebih baik daripada model A. Hal ini dipertegas lagi dari hasil CS pada model BB yang unggul pada batas error usia 5, 10, dan 25 tahun. Ditinjau

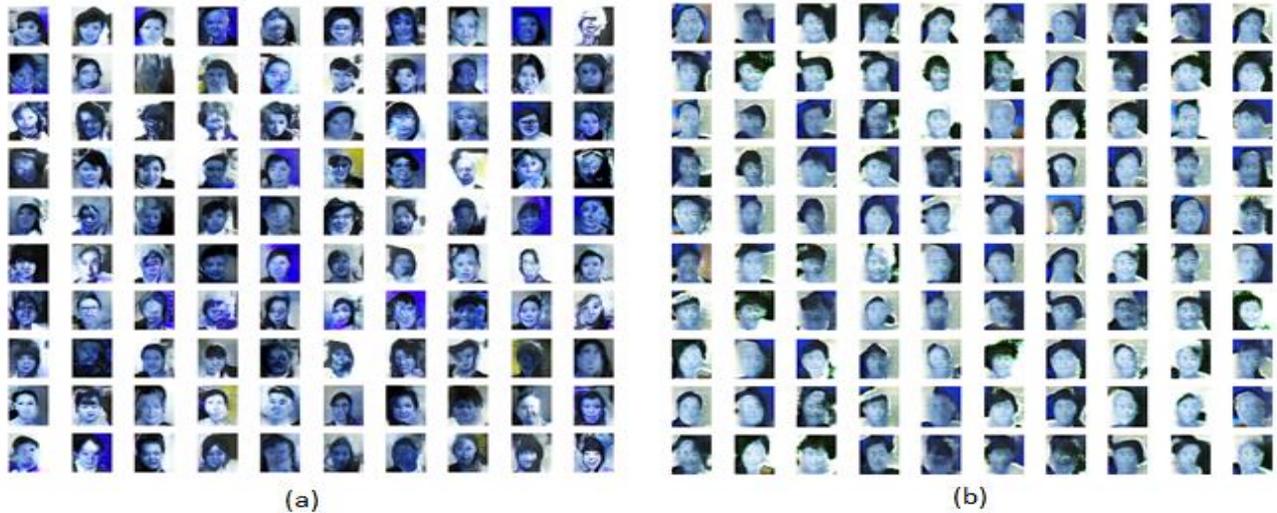
dari hasil batas error usia 5 dan 10 tahun ada kemungkinan 40-50% model mengalami kesalahan estimasi dengan batas error yang besar

Tabel 2 Hasil Skor MAE dan CS Data Testing

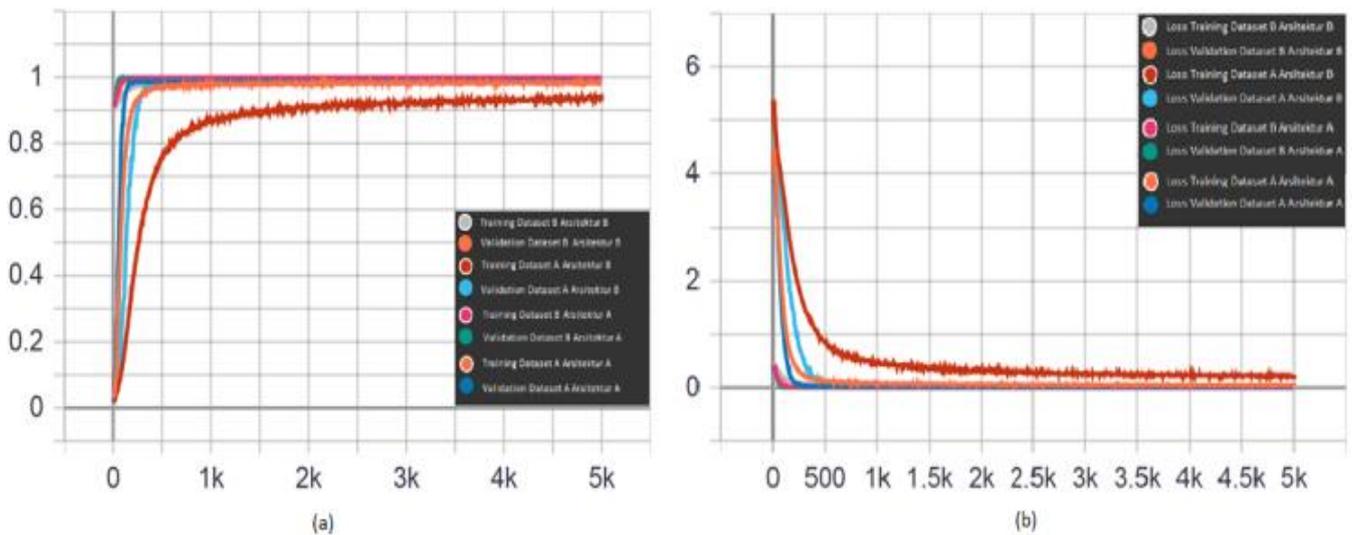
Model Age Estimation	MAE	CS ± SE
Model AA	14,195	51,1% ± 5,362
Model AB	12,964	53,6% ± 5,256
Model BA	14,144	51,5% ± 5,381
Model BB	12,847	56,3% ± 5,231

Tabel 3 Hasil Cumulative Score Data Testing Model AE Dari Beberapa Batas Error

Akurasi Usia Asli ± SE	Model AA	Model AB	Model BA	Model BB
CS(j) ≤ 5	32% ± 2,711	33,3% ± 2,546	31,4% ± 2,625	35,9% ± 2,659
CS(j) ≤ 10	51,1% ± 5,362	53,6% ± 5,256	51,5% ± 5,381	56,3% ± 5,231
CS(j) ≤ 15	65,5% ± 7,652	69,7% ± 7,709	66,1% ± 7,718	68,9% ± 7,240
CS(j) ≤ 20	75% ± 9,587	80% ± 9,656	75,7% ± 9,610	79,8% ± 9,421
CS(j) ≤ 25	82,5% ± 11,418	86,3% ± 11,154	83,1% ± 11,457	86,5% ± 11,068



Gambar 5 Perbandingan sampel gambar model generatif hyperparameter B latent space 150 pada (a) epoch ke-490 dan (b) epoch ke-500



Gambar 6 Hasil Training Model A dan Model B Age Estimation Menggunakan Dataset A dan Dataset B

C. Intelligent Advertising

Simulasi pada PC menggunakan empat jenis iklan untuk empat kategori usia yaitu anak-anak berusia 2-12 tahun, remaja berusia 13-24 tahun, dewasa berusia 25-59 tahun, dan lanjut usia (lansia) berusia 60-80 tahun. Karena sempitnya kategori usia maka iklan yang dapat ditampilkan sangat terbatas. Tampilan saat aplikasi bekerja dapat dilihat pada Gambar 7 yang menampilkan dua *windows*. *Window* pertama untuk menampilkan apa yang terlihat di kamera

seperti hasil estimasi usia dan lokasi wajah dari *face detection*. *Window* kedua untuk menampilkan iklan. Tampilan awal aplikasi sebelum bekerja menampilkan *window* tombol start yang akan berubah menjadi tampilan perekaman langsung dan *window* iklan dari usia acak sesuai kelas dataset yang ditentukan sebelum tampilan iklan muncul.

Aplikasi dilakukan pengujian terhadap audience langsung menggunakan model *age estimation* BB. Pemilihan model AE dilakukan sesuai hasil MAE dan

cumulative score terbaik. Estimasi menurut model BB mengalami error estimasi yang besar dan apabila berdiam cukup lama hasil estimasi langsung berubah bersama dengan iklan. Contohnya audience A berusia asli 76 tahun namun terestimasi sebagai usia 9 tahun dan setelah berdiam cukup lama hasil estimasi berubah menjadi 64 tahun. Penyesuaian iklan dapat ditampilkan dengan baik meskipun perubahan terhadap tingkat kecerahan atau ekspresi wajah dapat mengubah hasil estimasi dan iklan yang ditampilkan pada waktu yang sama.



Gambar 7 Aplikasi menampilkan hasil deteksi wajah dan estimasi usia serta mengubah *windows* iklan

V. KESIMPULAN

Berdasarkan pengujian fungsionalitas yang telah dilakukan oleh penulis, menunjukkan hasil bahwa aplikasi berfungsi dengan baik terlepas dari keterbatasan seperti variasi iklan dan tidak konsistennya estimasi usia akibat berbagai faktor. Hal ini dapat diketahui dari berhasilnya seluruh komponen aplikasi dan dapat disimpulkan bahwa penulis berhasil mensimulasikan *intelligent advertising* di Komputer.

Data augmentation menggunakan metode GAN berhasil meningkatkan performa hasil pada kedua arsitektur model AE terlepas dari rendahnya skor IS dan FID serta kualitas gambar yang jauh dari dataset asli. Hasil data augmentation lebih terlihat pada model B yang berhasil meningkatkan akurasi usia sebesar 4,8% dan skor MAE sebesar 1,297. Hasil MAE dari penggunaan dataset ini masih jauh dari apa yang sudah dibahas pada penelitian terkait. Selain itu, hasil pengukuran *cumulative score* batas error ≤ 25 yang mencapai 85% menunjukkan bahwa model AE memungkinkan adanya kesalahan estimasi yang sangat jauh dari usia asli.

Melihat dari hasil penelitian dan pengembangan yang telah dilakukan penulis, ada beberapa saran yang ingin disampaikan sehingga penelitian ini dapat dikembangkan kembali oleh peneliti selanjutnya seperti performa dari AE ini masih bisa ditingkatkan dengan memperkecil jumlah kelas dan menambah jumlah data terutama data usia lansia saat tahap data training. Perbedaan metode data augmentation juga memungkinkan adanya perbedaan performa dari model AE. Menambahkan fitur lain seperti gender detection dan object detection untuk mengambil lebih banyak data kemudian menambah variasi iklan yang dapat ditampilkan.

VI. DAFTAR PUSTAKA

- [1] G. M. Farinella, G. Farioli, S. Battiato, S. Leonardi, and S. Gallo, "Face Re-Identification For Digital Signage Applications," in *International Workshop on Video Analytics for Audience Measurement in Retail and Digital Signage*, 2014, vol. 8811, no. April 2016, pp. 40–52, doi: 10.1007/978-3-319-12811-5.
- [2] R. E. Abraham and M. R. Kennedy, "Intelligent digital signage system based on gender identification," in *Intelligent Embedded Systems*, 2018, vol. 492, pp. 251–262, doi: 10.1007/978-981-10-8575-8_25.
- [3] H. J. Müller and A. Krüger, "How much to bid in Digital Signage Advertising Auctions?," in *International Conference on Pervasive Computing (Pervasive 2007) Adjunct Workshop*, 2007, pp. 2–5.
- [4] T. J. Yu, C. P. Lee, K. M. Lim, and S. F. A. Razak, "AI-based targeted advertising system," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 2, pp. 787–793, 2019, doi: 10.11591/ijeecs.v13.i2.pp787-793.
- [5] A. S. Al-Shannaq and L. A. Elrefaie, "Comprehensive Analysis of the Literature for Age Estimation from Facial Images," *IEEE Access*, vol. 7, pp. 93229–93249, 2019, doi: 10.1109/ACCESS.2019.2927825.
- [6] X. Zhu, Y. Liu, J. Li, T. Wan, and Z. Qin, "Emotion classification with data augmentation using generative adversarial networks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10939 LNAI, pp. 349–360, 2018, doi: 10.1007/978-3-319-93040-4_28.
- [7] A. Antoniou, A. Storkey, and H. Edwards, "Data Augmentation Generative Adversarial Network," 2017, doi: 10.2139/ssrn.3349576.
- [8] S. Shao, P. Wang, and R. Yan, "Generative adversarial networks for data augmentation in machine fault diagnosis," *Comput. Ind.*, vol. 106, pp. 85–93, 2019, doi: 10.1016/j.compind.2019.01.001.
- [9] F.-Y. Cheng, C.-J. Chang, and G.-J. Jong, "Digital Interactive Kanban Advertisement System Using Face Recognition Methodology," *Comput. Water, Energy, Environ. Eng.*, vol. 02, no. 03, pp. 26–30, 2013, doi: 10.4236/cweee.2013.23b005.
- [10] E. Sümer, S. I. Sümer, and H. Oğul, "A novel computer vision-based advertisement system for individual marketing," *Teh. Vjesn.*, vol. 25, no. 5, pp. 1485–1491, 2018, doi: 10.17559/TV-20161220142243.
- [11] S.-H. Lee, M.-K. Sohn, and H. Kim, "Implementation of age and gender recognition system for intelligent digital signage," *Eighth Int. Conf. Mach. Vis. (ICMV 2015)*, vol. 9875, no. Icmv, p. 98750I, 2015, doi: 10.1117/12.2228520.
- [12] G. Levi and T. Hassner, "Age And Gender Classification Using Convolutional Neural Networks," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 34–42, doi: 10.1109/AFGR.2008.4813314.
- [13] F. Gurpınar, H. Kaya, H. Dibeklioglu, and A. A. Salah, "Kernel ELM and CNN Based Facial Age Estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2016, pp. 80–86, doi: 10.1109/CVPRW.2016.103.
- [14] S. H. Lee *et al.*, "Age and gender estimation using deep residual learning network," in *2018 International Workshop on Advanced Image Technology (IWAIT)*, 2018, pp. 1–3, doi: 10.1109/IWAIT.2018.8369763.
- [15] J. Cheng, Y. Li, J. Wang, L. Yu, and S. Wang, "Exploiting effective facial patches for robust gender recognition," *Tsinghua Sci. Technol.*, vol. 24, no. 3, pp. 333–345, 2019, doi: 10.26599/TST.2018.9010090.
- [16] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," *34th Int. Conf. Mach. Learn. ICML 2017*, vol. 6, pp. 4043–4055, 2017.