

**SURVEY TEKNIK PENGKLASIFIKASIAN GAYA  
ARSITEKTUR  
PADA FASAD BANGUNAN MENGGUNAKAN PENDEKATAN  
DEEP LEARNING CNN**

***SURVEY TECHNIQUE OF CLASSIFICATION OF  
ARCHITECTURAL STYLE ON BUILDING FACILITIES USING  
APPROACH DEEP LEARNING CNN***

Edy Sutomo  
Program Studi Teknik Arsitektur, Universitas Gunadarma  
*edysutomo@staff.gunadarma.ac.id*

**Abstrak**

Teknik pengklasifikasian gaya arsitektur pada fasad bangunan menjadi bagian penting pada dunia perancangan, guna mempercepat proses dalam melakukan kajian tipologinya. Dewasa ini dengan semakin berkembangnya teknologi informasi, sangat memungkinkan bila seiring waktu dengan berbagai kemajuan metode dalam mengekstraksi obyek bangunan utamanya fasad bangunan. Penelitian dalam pengklasifikasian fasad bangunan banyak dilakukan untuk menelusuri jenis bangunan maupun aspek estetika lainnya. Demi tujuan tersebut studi survey ini dimaksudkan untuk mengetahui teknik komputasi Deep Learning (DL) yang dapat digunakan dalam mengidentifikasi fasad bangunan secara lebih akurat dengan membedakan dan mengelompokkannya agar lebih mudah dikenali tipe bangunannya. Metode yang digunakan dalam melakukan penelitian ini menggunakan teknik seleksi dan eliminasi, berasal dari penelitian di berbagai jurnal yang relevan terhadap pengklasifikasian gaya arsitektur bangunan. Hasil survey literatur menunjukkan bahwa terdapat kesenjangan, hasil akurasi dari yang tertinggi ke terendah sebesar 48,19 % sehingga diperlukan adanya inovasi pada perangkat sistemnya. Teknik DL paling banyak digunakan dengan pendekatan Convolutional Neural Network (CNN) yang dikombinasikan dengan sistem perangkat lain, daripada fitur sendiri guna meningkatkan nilai akurasi.

**Kata kunci:** Gaya Arsitektur, Fasad, Deep Learning.

**Abstract**

*The technique of classifying architectural styles on building facades is an important part of the design world, in order to speed up the process of conducting typological studies. Nowadays, with the development of information technology, it is possible that over time with various advances in methods of extracting building objects, especially building facades. Research in the classification of building facades is carried out to explore the types of buildings and other aesthetic aspects. For this purpose, this survey study is intended to identify Deep Learning (DL) computational techniques that can be used to identify building facades more accurately by differentiating and grouping them to make it easier to identify the building type. The method used in conducting this research uses selection and elimination techniques, derived from research in various journals that are relevant to the classification of building architectural styles. The results of the literature survey show that there is a gap, the accuracy results from the highest to the lowest are 48,19% so that innovation is needed in the system equipment. The DL technique is most widely used with the Convolutional Neural Network (CNN) approach in combination with other device systems, rather than its own features to increase the accuracy value.*

**Keywords:** Architecture style, Facades, Deep Learning.

## PENDAHULUAN

Tipologi menurut kamus *webster* adalah studi atau analisis atau klasifikasi berdasarkan jenis atau kategori, dalam perjalannya studi tentang tipologi menjadi berkembang meliputi semua aspek dalam ilmu pengetahuan. Penelitian yang berhubungan dengan pengenalan bentuk visual tidak akan terlepas dari aspek-aspek yang berkaitan erat dengan teknik pengklasifikasian suatu obyek yang dapat dicerap oleh panca indera. Namun dalam perjalannya mengingat keterbatasan memori manusia sering kita jumpai bahwa kepastian dalam mengenali obyek visual dapat berubah-ubah atau tidak konsisten seiring kemampuan seseorang dalam mengingatnya. Demi tujuan tersebut maka arsitek dalam melaksanakan studi tipologi dapat menggunakan serta memanfaatkan teknologi komputasi yang semakin maju.

Pengklasifikasian merupakan salah satu aspek terpenting dalam dunia Arsitektur dalam rangka untuk menemukan atau melacak jenis atau tipe apakah yang digunakan dalam mendesain suatu bangunan. Fasad merupakan elemen utama dalam arsitektur yang dapat menyuarakan fungsi serta makna dari sebuah bangunan. Hal ini lebih mudah diidentifikasi melalui elemen penyusunannya, fasad secara implisit memberitahukan kondisi budaya ketika bangunan itu dibuat, fasad juga mengungkapkan parameter tatanan dan penataan, dan berjasa memberikan ruang kreatifitas dalam hal ornamentasi dan dekorasi. Kondisi fasad dari suatu gedung dapat menceritakan keadaan penghuninya serta merupakan wujud identitas kolektif dalam komunitas masyarakat tertentu (Sastra, 2016).

Sejarah pengklasifikasian gaya dan jenis arsitektur dapat ditinjau dari berbagai perspektif. Gaya arsitektur (mis., *Renaissance*, *Baroque*) menyediakan format dasar untuk mendesain bangunan individu di wilayah geografis pada zaman tertentu. Ornamen yang berkaitan dengan gaya tertentu dianggap

sebagai ekspresi keindahan untuk setiap zaman, fitur/ornamen ini kemudian mengubah bangunan biasa menjadi struktur dan arsitektur secara signifikan. Dengan demikian elemen visual seperti jendela, pilar, dan tatanan arsitektur dapat memberikan petunjuk untuk mengidentifikasi atau mengklasifikasikan arsitekturnya menjadi gaya tertentu.

Klasifikasi berbasis elemen pada arsitektur cagar budaya, sejumlah besar arsitektur modern dan kontemporer yang bertipologi fungsi, perubahan ini sebagian besar disebabkan oleh adanya perubahan konsep desain gaya Internasional yang bertujuan untuk mengekspresikan fungsi bangunan melalui "estetika mesin", menghasilkan pembentukan arsitektur modern menjadi *white box*. Ornamen dan dekorasi bersejarah ditolak dan "mesin" menjadi model arsitektur modern. Dengan demikian, kaum modernis cenderung mereduksi segala bentuk menjadi abstraksi. Selain itu, ruang dan pengalamannya menjadi salah satu topik terpenting dalam desain arsitektur modern dan kontemporer.

Kompleksitas pengklasifikasian karena ruang tidak dapat diidentifikasi oleh elemen, melainkan muncul ketika dikelilingi oleh kombinasi beberapa elemen spasial bersama dengan cahaya. DCNN secara teknis menurut cakupannya dapat berkontribusi terhadap klasifikasi gaya desain maupun klasifikasi yang berbasiskan elemen. Penggunaan teknik *Deep Learning* terbaru dapat dikembangkan dalam memproses gambar visual dalam mengklasifikasikan dataset yang diambil melalui sampel, kemudian dikelompokkan menurut kesamaan visual yang diukur oleh algoritma. Hasil akhir dapat dibandingkan dengan klasifikasi yang dibuat oleh sejarawan dan ahli teori arsitektur. Dengan demikian dapat ditunjukkan kemampuan kecerdasan buatan untuk mengklasifikasikan desain arsitektur modern dan kontemporer. Klasifikasi yang dihasilkan oleh komputer vision memberikan wawasan untuk mening-

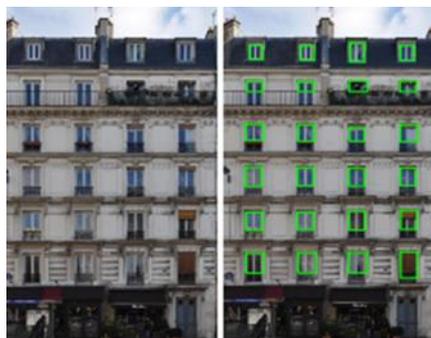
katkan pemahaman tentang desain arsitektur (Yoshimura et al., 2018).

Fasad disusun dari elemen tunggal, suatu kesatuan tersendiri yang mampu mengekspresikan diri mereka sendiri. Kata fasad berasal dari literatur bahasa Latin yaitu “*facies*” sinonim dari kata “*face*” (wajah) dan selanjutnya disebut “*appearance*” (penampilan). Fasad dapat dimaknai sebagai muka utama yang menghadap eksterior dari bangunan, berada pada sisi utama yang pada umumnya menghadap ke arah jalan utama. Fasad utama merupakan sisi bangunan yang mempunyai pintu masuk utama/entrance yang biasanya mempunyai detail yang menarik.(Sastra, 2016). Unsur pembentuk fasad meliputi pagar, ornamen dan semua atribut yang melekat pada muka bangunan.

Melalui tipologi suatu obyek arsitektur/bangunan bisa dianalisis/dijelaskan perubahannya yang terkait dengan bentuk/bangun dasar, sifat dasar, dan juga proses perkembangan maupun perubahan bangun dasar tersebut. Pengetahuan tipologi berikutnya akan berorientasi pada usaha dalam mengelaskan, mengklasifikasikan atau mengelompokkan melalui aspek atau aturan tertentu antara lain: (1) fungsi (terdiri dari struktural, simbolis, ruang dan lain sebagainya); (2) geometrik (terdiri dari prinsip tatanan, bentuk dan lain sebagainya); (3) langgam (terdiri dari periode, etnik, budaya, lokasi atau geografi, politik atau kekuasaan dan lain sebagainya). (Ramadanta,

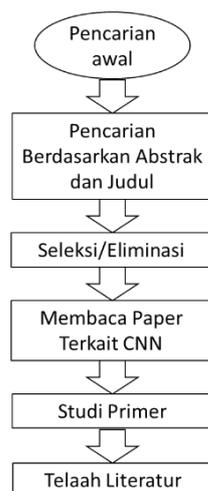
2010) Dalam kaitan ini langgam diusulkan menjadi *gaya/style* agar mudah dipahami untuk tujuan penelitian ini.

Pemanfaatan *Deep Learning* (DL) untuk *computer vision* (CV) memungkinkan sifat-sifat penglihatan manusia diterapkan pada komputer. Dapat berbentuk *smartphone*, *drone*, CCTV, pemindai MRI, dan sebagainya. Melalui berbagai sensor untuk penginderaan. Klasifikasi gambar merupakan pemberian label seluruh gambar dengan objek atau konsep. Aplikasi ini termasuk klasifikasi gender yang diambil dari gambar wajah seseorang, mengidentifikasi jenis hewan peliharaan, menandai foto, dan sebagainya. Keluaran klasifikasi dari fitur ini adalah deteksi, segmentasi, uji kemiripan, pelabelan gambar, generatif model dan analisis video. *Convolutional neural networks* (CNN) mirip dengan jaringan saraf yang di dalamnya memiliki bobot, bias, dan keluaran melalui aktivasi non linier. Jaringan saraf reguler mengambil input dan neuron yang terhubung sampai ke lapisan berikutnya. Neuron dalam lapisan yang sama tidak berbagi koneksi apa pun. Jika menggunakan jaringan saraf biasa untuk gambar, akan sangat besar ukurannya karena sejumlah besar neuron yang mengakibatkan *overfitting*. Tambah ukuran model (*resize model*) karena membutuhkan sejumlah besar *neuron*. Suatu gambar dapat dianggap sebagai volume dengan dimensi tinggi, lebar, dan kedalaman. (Sarkar et al., 2018)



**Gambar 1. Fasad Bangunan**

Sumber: Pesto,2016



**Gambar 2. Tahapan Survey Literatur**

Sumber: Penulis, 2020

Makalah ini bertujuan untuk melakukan survey penelitian tentang teknik pengklasifikasian fasad bangunan melalui teknik visi komputer, berdasarkan penampilan visual maupun elemen penyusunannya. Teknik *Deep Learning* CNN dapat mengidentifikasi fitur desain pada fasad bangunan dengan membedakan produk perancangannya dan mengelompokkannya menjadi tipologi arsitektur tertentu. Sehingga hasil akhir penelitian ini dapat menjadi referensi sebagai pengetahuan baru dalam bidang yang relevan.

## METODE PENELITIAN

Penelitian berfokus pada pencarian database ilmiah elektronik yang berasal dari berbagai pengelola jurnal ilmiah. Tahapan penelitian secara garis besar mengikuti kriteria inklusi antara lain terkait tentang studi pada penerapan CNN untuk teknik klasifikasi gaya arsitektur pada fasad bangunan dengan publikasi tahun 2010 sampai 2020, kemudian kriteria eliminasi berfokus pada studi yang berasal dari buku serta laporan penelitian ditahun yang sama.

Tahapan survey literatur dapat diuraikan sebagai berikut diawali dengan proses pencarian literatur serta proses yang mendasari CNN. Terdapat beberapa proses

dalam menyusun sebuah survey literatur, diantaranya yaitu proses pencarian awal, proses seleksi dan eliminasi, kemudian identifikasi studi primer dan yang terakhir telaah literatur. Mengingat keterbatasan pada penelitian maka dilakukan modifikasi dalam tahapan disesuaikan keperluannya. Secara umum proses survey literatur pada penelitian ini dapat dilihat pada gambar 2.

## HASIL DAN PEMBAHASAN

Penelitian ini bertujuan merekonstruksi pemodelan kota melalui pendekatan struktur buatan merubahnya menjadi label semantik secara otomatis, baik melalui gambar 2D maupun gambar 3D, selanjutnya klasifikasi dari dataset yang digunakan terdiri dari 4 kelas yaitu atap, dinding, bukaan (termasuk jendela dan pintu) dan balkon. Hasil menunjukkan bahwa fitur 3D dapat memperbaiki kesalahan dalam pengklasifikasian dengan piksel serta dapat memberikan informasi tambahan terkait spasial. Terdapat perbaikan kinerja dalam pengklasifikasian kelas balkon menggunakan FC-DenseNet yang terlatih dengan informasi 2D yang dikombinasikan dengan fitur 3D. Kinerja FC-DenseNet dan DeepLabV3 plus semuanya membaik dengan menambahkan fitur 3D. Di FC-DenseNet, IoU meningkat dari 59,28% menjadi 64,41%; Di DeepLab, IoU

meningkat dari 62,09% menjadi 62,16%. Akurasi keseluruhan juga semuanya meningkat dalam dua model: dari 88,42% menjadi 91,30% di FC-DenseNet dan dari 89,08% menjadi 91,10% masing-masing di DeepLab. FC-DenseNet yang terlatih dengan fitur 3D gabungan 3D mendapatkan hasil terbaik dengan akurasi 91,30% dan 64,41% IoU. Dalam beberapa tahapan terbukti bahwa fitur 3D dapat meningkatkan kinerja segmentasi, terjadi kebingungan membedakan antara atap dan dinding pada tahapan ini. Dibandingkan kedua model ini, DeepLabV3 plus memiliki kinerja yang lebih baik pada eksperimen 2D, dan FC-DenseNet mencapai kinerja terbaik dengan akurasi 91,30% dan 64,41% IoU saat menambahkan fitur 3D (Huang, 2019).

Dalam makalah ini menggunakan elemen jalur pipa untuk segmentasi semantik bangunan yang berasal dari gambar yang diperoleh melalui drone atau UAV (*Unmanned Aerial Vehicles*) dengan pendekatan CNN. Simulasi berbagai gambar tampilan udara dari gambar terestrial berdasarkan perkiraan homografi, melalui data pelatihan yang banyak tanpa memberi anotasi gambar UAV secara manual. Eksperimen ini menunjukkan bahwa jaringan terlatih dapat mencapai kinerja segmentasi yang wajar untuk gambar yang dihasilkan oleh UAV. Selain itu tampilan yang disimulasikan juga dapat ditingkatkan nilai akurasi segmentasinya secara signifikan. Gambar yang dihasilkan oleh UAV untuk pengujian dipilih data yang berasal dari dataset UAV yang tersedia, yaitu dataset ISPRS "Zeche Zollern" dan "Stadthaus", yang merupakan representasi bangunan bergaya modern dan tradisional. Kemudian dikumpulkan gambar fasad rumah yang berasal dari tempat yang terpisah di Morschenich, Jerman, menggunakan UAV rotari. Perbedaan dari dataset ISPRS di mana jendela ditampilkan dengan kaca transparan, beberapa jendela rumah ini ditutupi dengan tirai dan tampak buram. Untuk mengevaluasi

akurasi segmentasi, pelabelan gambar uji dilakukan secara manual digunakan sebagai data dasar. Hasil segmentasi piksel adalah sebagai berikut dengan perbandingan dari data asal meliputi obyek bangunan 84,92%, pintu 37,70%, jendela 80,39%, vegetasi 91,83% dan ground 83,11%. Data simulasi menunjukkan peningkatan nilai akurasi yaitu obyek bangunan 87,66%, pintu 41,87%, jendela 84,01%, vegetasi 92,60% dan ground 84,34% (Zhuo et al., 2019).

Obyek penelitian ini adalah pengklasifikasian gambar fasad bangunan yang memiliki elemen tangga dari jalan raya untuk dijadikan kelas utilitas yang berbeda, dengan tingkat kerumitan yang cukup tinggi sangat tergantung pada properti gambar, properti objek dan faktor lingkungan (variasi skala objek dan orientasi, perubahan jumlah bangunan per gambar, dipangkas fasad dalam gambar, oklusi, mengubah pencahayaan. Usulan penelitian ini melalui pendekatan *end to end* untuk mengklasifikasikan gambar tangga pada fasad bangunan melalui tingkat abstraksi yang tinggi hingga mencapai hasil yang sesuai untuk pengayaan model semantik perkotaan dalam 3D secara otomatis. Untuk memberikan penilaian yang wajar atas kinerja CNN. Selanjutnya melalui penggunaan dataset untuk mendapatkan data kinerja manusia sebagai perbandingan. Kebutuhan untuk studi seperti itu sangat jelas karena penampilan fasad tidak selalu sesuai dengan penggunaan aktual bangunan atau alih fungsi. Asumsi bahwa perbedaan ini tidak dapat diselesaikan dalam ruang visual, maka dengan adanya batas alami untuk mendapatkan nilai akurasi secara keseluruhan. Demikian tingkat kesalahan klasifikasi gambar 36% (kasus 5 kelas) diperlukan perbaikan lanjutan. Penelitian ini mencoba dengan merubah metode pendekatan, melalui peningkatan partisi kelas/klasifikasi dan mengurangi variansi intra kelas. Perlunya penggunaan model mutakhir (VGG16, VGG19, ResNet50) karena cukup kompleks tingkat kesulitannya, maka hasilnya kinerja

sistemnya menjadi sedikit lebih buruk. Perbaikannya digunakan data pratraining yang berasal dari ImageNet. Hasil data pratraining meningkatkan kinerja sebesar 8-10% meskipun gambar yang berasal dari ImageNet sangat berbeda dengan dataset. Jaringan yang dirancang sendiri dapat mencapai akurasi secara keseluruhan 52,29% sekaligus mengurangi jumlah parameter sekitar 72% (Laupheimer & Haala, 2018).

Kontribusi penelitian ini bermaksud untuk penyediaan set data pelatihan dengan menghubungkan informasi semantik yang sudah ada pada database yang disediakan dari badan pemetaan nasional atau administrasi kota ke gambar fasad bangunan terkait gambar yang diambil dari SV (*street view*). Konsep penelitian ini menggunakan CNN dan classifier sebagai alternatif. Data untuk pelatihan dan pengujian CNN yang sama dengan sebelumnya, jumlah dataset pelatihan terdiri dari 8000 gambar (4 kelas, masing-masing 2.000 gambar) dan dataset validasi berisi berjumlah 70 gambar per kelas. Adapun data latih sejumlah 2.200 gambar dengan distribusi untuk bangunan komersial 19%, hibrid/campuran 22%, perumahan 43% dan penggunaan khusus 16%. Data simulasi menunjukkan hasil peningkatan nilai akurasi berdasarkan klasifikasi fungsi bangunan sebelumnya recall yaitu obyek komersial 71,62% campuran/hibrid 76,80%, perumahan 75,89% dan peruntukan khusus 78,48 %. Setelah dievaluasi akurasi dapat ditingkatkan untuk obyek komersial 72,60%, campuran/hibrid 80,67%, perumahan 85,00% dan peruntukan khusus 63,27% (Tutzauer & Haala, 2017).

Studi ini adalah membuat dataset baru dengan cara memasukkan gambar struktur asitektonik bangunan yang berasal dari Mexico terbagi menjadi 3 kategori yaitu prahispanik, kolonial dan modern. Gambar berasal dari lembaga budaya Meksiko berupa video yang diekstraksi. Dataset berisi berbagai jenis gambar dalam format perspektif

bangunan yang di dalamnya termasuk keempat kategori dalam klasifikasi tersebut dan gambar non-arsitektur. Agar dapat meningkatkan jumlah gambar dalam dataset dapat diperoleh melalui transformasi. Kemudian matriks yang tidak konsisten (*class confusion*) kelas dari pelatihan dengan kinerja terbaik, hal ini sangat berguna untuk mendeteksi evolusi pembelajaran di kelas selama fase pelatihan. Pada diagonal utama, persentase prediksi yang benar dari setiap klasifikasi dapat dilihat secara individual. Setiap klasifikasi berupa besaran prosentase ketidakkonsistenan dengan yang lain disebabkan adanya kesamaan antar kelas pada gambar yang terdapat pada database. Dalam gambar *prehispanic*, terdapat banyak konten alam, langit, atau tumbuhan seperti di klasifikasi lain dan sebaliknya terdapat konten sejenis pada gambar kelas lain namun tanpa arsitektur *prehispanic*. Hal ini menyebabkan kebingungan dengan kelas lain, perilaku ideal pengklasifikasian harus terdapat diskriminasi gaya arsitektur dari kelas lain. Hasil ringkasan perbandingan antara dataset dengan label prediksi dengan 4 kelas untuk data set yaitu Prahispanik 89,16%, Kolonial 87,95%, Modern 83,39% dan lainnya 91,96% (Obeso et al., 2016).

Penelitian ini bertujuan membuat klasifikasi gambar dan pemulihan gambar berbasis konten (CBIR) menggunakan *convolutional neural networks* (CNN). Dalam hal ini dengan mengusulkan representasi tingkat menengah berupa pembaharuan/novelty menggunakan CNN pra terlatih, untuk mengekstraksi fitur dan menggunakannya dalam menyelesaikan pengklasifikasian, serta pengambilan pada dataset gambar bangunan dengan gaya arsitektur yang berbeda. Dataset berupa gambar warna bangunan dari 25 gaya arsitektur yang berbeda berisikan 4.794 foto. Gambar ini diunduh dari koleksi Wikimedia dan menampilkan banyak pilihan dari era yang berbeda. Klasifikasi terdiri dari kelas gaya arsitektur Mesir kuno, *American craftsman*,

*American square styles*, Romawi, Gotik dan selebihnya adalah gaya campuran. Dalam hal ini obyek penelitiannya adalah komparasi metode lain dengan CNN yang sudah dimodifikasi dengan menggunakan 25 kelas gaya arsitektur. Hasilnya dapat dijelaskan sebagai berikut GIST 17.39%, SP 44.52%, OB-Partless 42.50%, OB-Part 45.41%, DPM-LSVM 37.69%, DPM-MLLR 42.55%, MLLR+SP 46.21% dan diperbaiki dengan CNN layer 16 Raw+EFM 63.90% (Meltser et al., 2018).

Pengklasifikasian gambar dalam penelitian ini menggunakan *Convolutional Neural Networks* (CNN) dan segmentasi gambar berdasarkan piksel, menunjukkan fleksibilitas ketika mendeteksi dan mengklasifikasikan berbagai jenis objek, bahkan dalam situasi di mana manusia gagal memahami perbedaan dalam skenario kota. Penelitian ini dilakukan dengan cara mengeksplorasi dan bereksperimen dengan teknik pelabelan semantik pada model perkotaan dalam format 3D. Tahapan penelitian dibagi menjadi dua jalur pemrosesan utama: pertama, cara memberi label pada fitur fasad dalam domain 2D, di mana CNN *supervised* digunakan untuk mengelompokkan gambar fasad berbasis darat menjadi enam kelas fitur yaitu, atap, jendela, dinding, pintu, balkon dan toko; kedua, alur kerja *Structure-from-Motion* (SfM) dan *Multi-View-Stereo* (MVS) digunakan untuk mengekstraksi elemen geometri pada fasad bangunan, di mana gambar tersegmentasi pada tahap sebelumnya kemudian digunakan label mesh yang dihasilkan dari "inverse" teknik penelusuran sinar. Kumpulan data SJC merupakan bangunan perumahan di São José dos Campos, São Paulo, Brasil. Gaya arsitektur yang terdapat di seluruh kota ini tidak unik (campuran), sering menyimpang atau tidak jelas antara gaya bebas dan gaya modern. Dataset sebanyak 175 gambar secara berurutan serta tumpang tindih diambil pada saat yang sama. Makalah ini menunjukkan

bahwa metodologi yang diusulkan cukup memadai dalam skenario yang kompleks. Kesimpulan fitur fasad telah mencapai akurasi hingga 93% atas sebagian besar dataset yang digunakan (Lotte et al., 2018).

Dalam penelitian ini dilakukan eksplorasi dalam pembelajaran mengkategorikan gambar ke dalam kelas abstrak dimungkinkan karena tidak adanya batas visual yang jelas antara kategori dan gambar, misalnya mengidentifikasi citra dengan suasana hati senang, sedih dan netral. Algoritma diusulkan melalui menggabungkan CNN dan prior ontologi untuk menyimpulkan pola abstrak dalam gambar menara pada bangunan yang menyerupai masjid di India. Pendekatan yang digunakan berbasis pembelajaran transfer di mana, pengetahuan (*knowledge*) sebagai domain ditransfer ke CNN saat pelatihan (*transfer top-down*) dan inferensi dibuat menggunakan prediksi CNN serta pohon ontologi / prior (*transfer bottom-up*). Kemudian dengan mengklasifikasikan gambar ke dalam kategori seperti Makam, Benteng dan Masjid. Melalui pendekatan metode tersebut terjadi peningkatan yang signifikan dibandingkan dengan pengelompokan logistik menggunakan pendekatan pembelajaran transfer lainnya. Hasilnya dengan menggunakan 3 *classifier* yaitu Logistik 54.62%, Obyektif 1 63.55%, Obyektif 2 71.43% diperbaiki melalui pendekatan *leaf node accuracy* menjadi Obyektif 1 21.51%, Obyektif 2 75.59% (Gupta & Chaudhury, 2016).

Penelitian ini mengeksplorasi algoritma yang menyediakan fasilitas klasifikasi piksel untuk fasad bangunan, dikarenakan fasad bangunan terdapat lingkungan yang kaya elemen untuk menguji teknik segmentasi semantik. Fasad bangunan dengan berbagai gaya akan mencerminkan karakteristik visual dan tata letak, selain itu juga menunjukkan tingkat stabilitas dalam pengaturan pola struktur material. Alasan tersebut menjadikan sebab untuk mengintegrasikan tampilan visual

dan petunjuk tata letak dalam satu kerangka. Pelabelan yang paling memungkinkan yaitu berdasarkan penampilan visual diperoleh melalui penerapan *Convolutional Neural Networks* (CNN) yang telah dimodifikasi dan dioptimasi dengan *Restricted Boltzmann Machines* (RBM) selanjutnya diterapkan pada scanlines vertikal dan horizontal dari fasad bangunan. Tahapan berikutnya menguji algoritma pada dataset ECP-Monge dan dataset CMP. ECP-Monge berisi 104 gambar yang diperbaiki dari fasad dalam gaya Hausmannian. Terdapat 8 struktur yang ditentukan dalam peta *groundtruth*. Dataset CMP berisi sejumlah 378 sampel dengan 12 struktur dari berbagai gaya (seringkali sulit untuk model). Selanjutnya dalam percobaan, sy dan sx disatukan pada semua gambar serta diatur masing-masing menjadi 300 dan 200. Hasil untuk dataset ECPMonge 90,49% dan dataset CMP adalah 65,54% (Fathalla & Vogiatzis, 2017).

Pemetaan gaya arsitektur terhadap rumah dari blok perumahan di Amerika Serikat yang memiliki gaya arsitektur khas daerah diantaranya gaya Ranch dan Craftsman. Obyek yang diteliti meliputi bagian fasad bangunan termasuk elemen unsur pembentuknya. Dataset model yang digunakan berasal dari ResNet-18 dan ResNet-34, terdiri dari 2500 gambar yang didistribusikan ke dalam 5 kelas sehingga masing-masing menggunakan 500 gambar. Kelas dibuat berdasarkan gaya arsitektur yang berbeda pada wilayah perumahan tersebut. Metode pengklasifikasian menggunakan kerangka *Deep Learning* Pytorch untuk model *baseline* pelatihan *end to end* dan bobot model pratraining. Hasil menunjukkan bahwa sangat memungkinkan untuk mengklasifikasikan gaya arsitektur rumah di AS. dengan CNN. Akurasi mencapai tingkat klasifikasi hingga 79,8%, dan nilai skor lokalisasi deviasi 0,710 pada dataset tes dengan menggunakan ResNet-34 sebagai ekstraktor fitur (Pesto, 2016).

Ekstraksi fasad bangunan merupakan

komponen kunci dalam merekonstruksi *street scene* dalam bentuk 3D. Penelitian ini mengusulkan metode *Deep Learning* untuk segmentasi fasad ke dalam klasifikasi semantik. Struktur buatan manusia sering menghadirkan karakteristik yang simetri, terkait pada masalah ini mengusulkan adanya regulator/pengatur simetris untuk melatih CNN. Metode yang diusulkan dapat memanfaatkan kekuatan jaringan saraf yang terdapat pada struktur arsitektur buatan manusia, melalui perbaikan hasil segmentasi dengan menggunakan kotak partisi yang dihasilkan oleh *Region Proposal Network* melalui cara melatih jaringan FCN-8 dengan *function loss* baru. Hasil eksperimen menunjukkan bahwa metode yang diusulkan melampaui metode sebelumnya secara signifikan pada dataset ECP maupun dataset eTRIMS. Aturan buatan dimasukkan ke dalam sistem *end-to-end* dari *Deep Convolutional Neural Networks* (DCNN), perlunya penggunaan istilah baru yaitu kerugian (*loss*) berdasarkan simetri yang biasanya ditemukan dalam struktur seperti bukaan pada jendela, dinding, dan pintu. Teknik segmentasi juga ditemukan deteksi berbasis *Deep Learning* (DL) sangat membantu dalam prosedur parsing di samping segmentasi pipa ujung ke ujung bangunan. Hasil dengan menggunakan klasifikasi terdiri dari Bangunan 96.03%, Mobil 94.20%, Pintu 80.66%, Perkerasan 84.81%, Jalan 90.58%, Langit 98.06%, Vegetasi 94.16%, Jendela 90.91%, dan total akumulasi 91.18% (Liu et al., 2017).

Parsing fasad menjadi masalah mendasar dalam pemodelan perkotaan yang merupakan tulang punggung dari masalah pemodelan berbasis prosedur, analisis arsitektur, dan rekonstruksi perkotaan, hal ini sangat bergantung pada segmentasi semantik. Dengan beralih ke pendekatan berbasis *Deep Learning*, dataset skala kecil yang ada menjadi hambatan untuk membuat kemajuan lebih lanjut dalam segmentasi fasad yang berdampak pada fasad parsing. Dalam tesis ini

mengusulkan dataset gambar fasad baru dalam segmentasi semantik yang disebut sebagai PSV-22, yang merupakan dataset terbesar serta memungkinkan adanya perbaikan hasil dari semantik fasad dari dataset yang ada. Di samping itu melalui tiga modifikasi segmentasi semantik berbasis *Deep Learning* menunjukkan bahwa modifikasi ini meningkatkan kinerja pada dataset perbaikan dan dataset yang sudah ada. Pada bagian ini dengan mengevaluasi dataset PSV baru untuk menunjukkan bahwa hal itu mengarah ke model yang lebih umum. Selain itu dengan menggunakan dataset eTRIMS, ECP, dan CMP, guna membandingkan teknik pendekatan sebelumnya pada set data yang telah ditetapkan. Dengan menggunakan kombinasi data PSV dan CMP untuk studi ablasi karena menggunakan serangkaian label yang umum, sehingga hasilnya yang didasarkan pada 1200 gambar yang dipilih acak. Secara kuantitatif untuk pelabelan pada PSV (dikombinasikan dengan CMP) data menggunakan DLV3 + MES. Hasil penelitian menunjukkan akurasi dengan klasifikasi fasad berdasarkan elemen berikut: Fasad 95%, Jendela 93%, Pintu 79%, Cornice 87%, Ambang 81%, Balkon 89%, Blind 90%, dekorasi 69%, kolom 92%, Toko 88% dan kumulatif rerata 96,74% (Para, 2019).

Dalam tulisan ini mengusulkan pendekatan *end-to-end* untuk mengklasifikasikan gambar terestrial dari fasad bangunan menjadi lima kelas utilitas yang berbeda (komersial, hibrida, perumahan, penggunaan khusus, dalam konstruksi) dengan menggunakan *Convolutional Neural Networks* (CNN). Dengan menggunakan gambar yang disediakan oleh *Google Street View* gambar tersebut secara otomatis ditautkan ke model kota kasar, termasuk garis sempadan bangunan serta kelas penggunaan masing-masing sesuai fungsinya. Dengan dataset yang tersedia untuk pelatihan dan evaluasi pipa menggunakan *Deep Learning*. Makalah ini menjelaskan pendekatan ujung ke ujung yang

diimplementasikan untuk mengklasifikasikan gambar jalan tangga dari fasad bangunan serta membahas eksperimen dengan berbagai metode pendekatan CNN. Selain hasil klasifikasi, yang disebut *Class Activation Maps* (CAM) dievaluasi. Kemudian peta akan menunjukkan bagian fasad yang telah ditentukan serta digunakan sebagai fitur selama proses pelatihan. Selanjutnya peta dapat digunakan untuk meneruskan presentasi abstrak dengan difasilitasi pemahaman berkonten gambar semantik. Representasi yang abstrak merupakan hasil dari metode penetapan, rendering gambar berbasis kegunaan/fungsi. Melalui bantuan dengan penambahan data priori (*flipping horizontal, warping, cropping, jittering*, dan modifikasi *saturation*) dengan membuat dataset yang terdistribusi secara merata menggunakan jumlah sampel yang sama untuk setiap kelas. Dataset pelatihan terakhir menggunakan dataset yang seimbang terdiri dari 75.000 gambar berlabel secara total (15.000 gambar per kelas/klasifikasi). Perangkat validasi dan perangkat uji masing-masing terdiri dari 350 gambar berlabel (70 gambar per kelas). Ukuran uji dan validasi yang relatif kecil ini disebabkan oleh pemisahan dataset sebelum augmentasi data. Semua paket yang disediakan dibagi menjadi tiga bagian: gambar untuk pelatihan (*training set*), gambar untuk memvalidasi kinerja jaringan selama proses pelatihan (*validation set*) dan gambar untuk mengevaluasi kinerja jaringan setelah proses pelatihan (*sample set*). Hasilnya akhir sebagai berikut yaitu VGG 16 63.43%, VGG19 63.14%, ResNet 50 55.43% dan Interception V3 64.00% rerata sebesar 60,25 % (D. Laupheimer et al., 2018)

Makalah ini meneliti guna mengenali elemen berbagai bangunan untuk mengambil serta mengenali spesifikasi teknis dari literatur perdagangan kontemporer. Metode mengambil studi kasus Perpustakaan Umum Carnegie di Inggris. Dalam hal probabilitas Pengenalan obyek mencapai nilai di atas ambang batas

arbitrer, dengan kotak terikat secara spesifik dianggap mengandung obyek yang sesuai. Area algoritma menghasilkan ribuan kotak yang berbeda, prosesnya terikat sehingga menghasilkan deteksi yang tumpang. Karena itu, sebagai langkah terakhir, penumpukan tidak maksimal harus diterapkan untuk membersihkan deteksi. Kinerja *benchmark Traditional Machine Learning* (TML), *Convolutional Neural Networks* (CNN) dan *Feature Pyramid Neural* (FPN) dapat dilihat pada bagian akhir paragraf. Jika kita melatih ulang perbandingannya dimungkinkan untuk melihat bahwa arsitektur FPN yang lebih maju tampil, seperti yang diharapkan, sedikit lebih baik daripada arsitektur CNN klasik. Hasil akhir akurasi perbandingan dengan menggunakan 5 klasifikasi adalah sebagai berikut: TML 56%, CNN 78%, dan FPN 82% (Pezzica et al., 2019).

Penelitian ini menerapkan teknik *Deep Learning* (DL) dan visi komputer untuk mengukur kesamaan visual antara desain arsitektur bangunan oleh arsitek yang berbeda. Melalui pemeriksaan bobot dalam model DCNN yang terlatih, dapat diukur secara kuantitatif kesamaan visual antara karya arsitek yang secara implisit dipelajari oleh model DCNN. Penggunaan ukuran ini melalui mengelompokkan arsitek yang diidentifikasi serupa dan membandingkan hasil temuan dengan klasifikasi konvensional yang dibuat oleh sejarawan dan ahli teori arsitektur. Pengelompokan desain arsitektur bangunan dengan hasil sangat menguatkan pandangan konvensional dalam sejarah arsitektur, dan fitur arsitektur yang dipelajari juga sejalan dengan pemahaman tradisional tentang DL. Dengan menggunakan dataset yang terdiri dari gambar yang dikikis dari web serta koleksi yang asli yang berasal dari gambar karya arsitektur sejumlah 19.568 bersumber dari berbagai web, pertama dengan melatih model *neural network* (DCNN) konvensional yang mampu mencapai akurasi 73% dalam mengklasifikasikan karya-karya milik 34

arsitek yang berbeda (Yoshimura et al., 2018)

Riset ini merupakan laporan akhir proyek yang dilakukan pada tahun terakhir sarjana Ilmu Komputer. Informasi ini terkait tentang masalah yang diusulkan oleh De Energie bespaarders, sebuah perusahaan start-up di Amsterdam yang berkonsentrasi pada penghematan energi. Laporan tersebut menjelaskan gambar rumah dengan pendekatan penerapan pembelajaran mesin (DL) dalam bentuk jaringan saraf (CNN) untuk mengenali jendela, pintu, dan dinding. Tujuan awal dari jaringan saraf adalah pengklasifikasian setiap fitur yang ditemukan melalui segmentasi berupa bukaan jendela atau pintu. Klasifikasi akan terdokumentasi dengan baik menggunakan jaringan saraf konvolusional (CNN), mencapai akurasi setinggi 93,72% untuk kumpulan data standar seperti CIFAR-10 (Kolenbrander et al., 2017)

Melalui media *Google Earth imagery* terdapat 1005 gambar spasial yang terbagi menjadi 19 kelas beresolusi tinggi dengan  $600 \times 600$  piksel. Terdapat setidaknya 50 gambar mewakili setiap kelas yang diperoleh melalui pemandangan udara. Sejauh ini, dataset WHU-RS digunakan secara ekstensif untuk penelitian metode klasifikasi yang berasal dari pemandangan udara. Seting percobaan melalui simulasi dengan pengklasifikasian berdasarkan karakteristik yang berasal dari lapisan jaringan kerja (*networks*). Hasil yang diperoleh melalui 80% dataset UC-Merced digunakan untuk set pelatihan dan dataset WHU-RS dengan rasio pelatihan 60%. (Petrovska et al., 2020)

NetVLAD merupakan lapisan VLAD terinspirasi oleh tampilan "*Vector of Locally Aggregated Descriptors*" yang digunakan dalam pengambilan gambar. Layer tersebut dihubungkan ke dalam arsitektur CNN dengan pelatihan melalui propagasi mundur. Terjadi penurunan kinerja peringkat dari *unsupervised* dataset, memungkinkan pembelajaran menyeluruh dari parameter arsitektur dari gambar, tempat yang sama pada saat diunduh

dari *Time Machine Google Street View*. Melalui Layer *Full Convolutional (FC)* yang terhubung penuh terdapat beberapa masalah yaitu: (i) tidak dapat berubah atau tidak dapat diubah ukuran layernya, (ii) ukurannya menjadi sangat besar pada fokus utamanya, (iii) terjemahan menjadi tidak invarian pada skala atau oklusi parsial karena memaksakan ketetapan tata letak spasial. Eksperimen di Tokyo diketahui bahwa penggabungan diperlukan, karena *baseline "no pool"* yang terlatih capaian terbaik pada *recall @ 5* dari 63,5%, berbanding dengan 82,5% pada NetVLAD.. (Arandjelovic et al., 2018)

Fokus penelitian pada bangunan di perkotaan dengan mengklasifikasikan menjadi 10 kategori yaitu gereja, masjid, sinagoge, wihara, apartemen, mal, gudang, restoran, dan bangunan perkantoran menggunakan multi hirarki label CNN yaitu *Branch Convolutional Neural Network (B-CNN)*. Menggunakan 2 dataset dengan gaya arsitektur yang berbeda untuk 284 gambar, selanjutnya terdapat 4 kelas kategori dengan 2000 gambar namun hanya 1028 yang dapat dikenali dengan gaya arsitekturnya, berikutnya ditingkatkan

menggunakan gaya arsitektur dengan 6 kelas meliputi Gotik, Bisantium, Baroq, Modern, Islam dan kolonial akurasinya 92,65 % menggunakan 21.190.606 jenis parameter, total Loss 0,89. (Taoufiq et al., 2020)

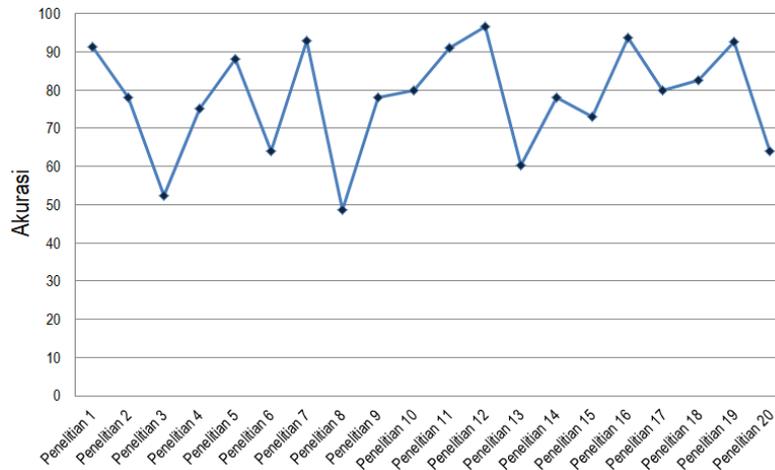
CBIR merupakan teknik klasifikasi dan pengambilan gambar berbasis konten di mana terdapat permasalahan kompleks pada bidang visi komputernya. Jaringan saraf konvolusional (CNN) menjadi pilihan untuk membangun sistem pengklasifikasian gambar. Dataset ini terdiri dari gambar bangunan berwarna terdapat 25 gaya arsitektur berbeda, yang berisi 4794 foto. Gambar diunduh dari Wikimedia dengan menampilkan banyak pilihan dari era yang berbeda. Nama kategori berikut jumlah gambar masing-masing dataset. Percobaan menggunakan metode kation klasifikasi dengan menggunakan 30 gambar dari setiap kelas untuk pelatihan klasifikasi. Fitur CNN pada lapisan perantara (lapisan 9-16) mengungguli lapisan bawah dan atas. Akurasi klasifikasi tertinggi yang didapatkan dengan klasifikasi gaya arsitektur adalah 63,9% lebih besar 17% peningkatannya dari metode MLLR + SP. (Meltser et al., 2018)

**Tabel 1. Deep Learning Untuk Jenis Kategori dan Akurasi**

| Rujukan       | Jenis obyek data | Jumlah Kategori | Akurasi |
|---------------|------------------|-----------------|---------|
| Penelitian 1  | Gambar           | 4               | 91,30 % |
| Penelitian 2  | Gambar           | 10              | 78,10 % |
| Penelitian 3  | Gambar           | 5               | 52,29 % |
| Penelitian 4  | Gambar           | 4               | 75,14 % |
| Penelitian 5  | Video            | 4               | 88,12 % |
| Penelitian 6  | Gambar           | 25              | 63,90 % |
| Penelitian 7  | Gambar           | 3               | 93,00 % |
| Penelitian 8  | Gambar           | 3               | 48,55 % |
| Penelitian 9  | Gambar           | 12              | 78,02 % |
| Penelitian 10 | Gambar           | 5               | 79,80 % |
| Penelitian 11 | Gambar           | 8               | 91,18 % |
| Penelitian 12 | Gambar           | 10              | 96,74 % |
| Penelitian 13 | Gambar           | 5               | 60,25 % |
| Penelitian 14 | Gambar           | 5               | 78,00%  |
| Penelitian 15 | Gambar           | 34              | 73,00 % |
| Penelitian 16 | Gambar           | 3               | 93,72 % |

|               |        |    |         |
|---------------|--------|----|---------|
| Penelitian 17 | Gambar | 19 | 80,00 % |
| Penelitian 18 | Gambar | -  | 82,50 % |
| Penelitian 19 | Gambar | 6  | 92,65 % |
| Penelitian 20 | Gambar | 25 | 63,90 % |

Sumber: Penulis,2021



**Grafik 1. Deep Learning untuk Akurasi**

Sumber: Penulis,2021

Setelah meninjau berbagai paper di atas dapat dijelaskan sebagai berikut bahwa penggunaan *Convolutional Neural Network* (CNN) banyak digunakan secara murni dengan semua fitur yang terdapat pada metode tersebut, namun pada beberapa penelitian CNN dimodifikasi sesuai kebutuhan, diantaranya dengan mengembangkan algoritma serta *Restricted Boltzmann Machines* (RBM). Fenomena yang dapat dipahami dalam survey literatur ini adalah adanya keinginan untuk meningkatkan kinerja dari CNN dalam melakukan pengklasifikasian

fasad bangunan. Hasil dari masing-masing penelitian tingkat akurasinya paling rendah yaitu penelitian 8 sebesar 48.55% dan yang paling tinggi nilai akurasinya dilakukan oleh penelitian 12 sebesar 96,74%. Komparasi hasil akurasi tiap penelitian dapat dilihat pada grafik 1 dan Tabel 1. Nilai akurasi yang rendah tersebut dikarenakan metode deteksi gambar sebanyak 1200 data. Sebaliknya nilai akurasi dapat capai dengan prosentase yang besar salah satunya disebabkan segmentasi fasad menggunakan klasifikasi dalam *framework* semantik.

**Tabel 2. Deep Learning Untuk Teknik Pengklasifikasian**

| Teknik Pendekatan<br>Terkait      | Jumlah Paper |
|-----------------------------------|--------------|
| Klasifikasi gambar                | 20 paper     |
| Deteksi,lokalisasi,<br>Segmentasi | 20 paper     |
| Kemiripan                         | 20 paper     |
| Diskripsi gambar                  | 3 paper      |
| Generatif model                   | 2 paper      |
| Analisis video                    | 1 paper      |

Sumber: Penulis,2021

Pada tabel 2 berisikan ringkasan yang memuat tentang teknik pendekatan yang digunakan terhadap obyek risetnya. Hal ini didapatkan melalui hasil studi dari *state of the arts* pada penelitian sebelumnya, terkait penggunaan *Deep Learning Convolutional Neural Networks* (DLCNN) baik yang murni dari fitur aslinya, maupun dengan modifikasi dalam sistem perangkatnya.

Berdasarkan tabel 2 dalam melakukan penelitian menggunakan pendekatan *deep learning CNN* menunjukkan trend bahwa klasifikasi gambar menjadi pilihan paling banyak digunakan dalam melakukan analisis terhadap obyek gambar. Terbukti bahwa teknik pendeteksian digunakan untuk klasifikasi gambar digunakan dalam 20 paper, deteksi lokasi dan segmentasi dipakai dalam penelitian sebanyak 20 paper, berikutnya kemiripan juga digunakan dalam penelitian terdapat pada paper sebanyak 20 paper, selanjutnya diskripsi gambar digunakan oleh peneliti dalam 3 paper, berikutnya fitur generatif model digunakan penelitian dalam 2 paper dan yang terakhir adalah fitur analisis video hanya 1 yang menggunakannya. Hal ini dikarenakan peneliti pada umumnya menggunakan pendekatan yang memiliki akurasi tinggi namun tidak terlalu kompleks peralatannya.

## SIMPULAN

Secara umum penelitian yang disurvei pada penggunaan *Deep Learning* dengan metode *Convolutional Neural Networks*, dalam melakukan pengklasifikasian gaya arsitektur pada fasad bangunan dapat memberikan hasil yang cukup memuaskan, didukung dengan dataset training yang memadai. Hasil akurasi yang rendah dan sangat tinggi terjadi dikarenakan beberapa faktor antara lain terlalu banyaknya dataset yang ditraining, fungsi kerugian (*loss function*), optimisasi, regularisasi, algoritma pembelajaran, dan inovasi dalam perangkat sistemnya.

Penerapan *Deep Learning* dengan metode *Convolutional Neural Networks* di dalam klasifikasi pada fasad bangunan menunjukkan terdapat bermacam variasi baik berupa teknik murni, maupun dengan cara memodifikasi untuk berbagai tujuan dan fungsi. Kontribusi rekayasa tersebut juga tersebar ke dalam berbagai bidang penelitian. Penelitian ini merekomendasikan beberapa arah penelitian yang dipandang perlu untuk diteliti dimasa depan dalam lingkup *Deep Learning* dengan metode *Convolutional Neural Networks* dalam proses perancangan Arsitektur bangunan.

## DAFTAR PUSTAKA

- Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2018). *NetVLAD: CNN Architecture for Weakly Supervised Place Recognition*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6), 1437–1451. <https://doi.org/10.1109/TPAMI.2017.2711011> [viewed on 01/02/2021]
- Fathalla, R., & Vogiatzis, G. (2017). *A deep learning pipeline for semantic facade segmentation*. *British Machine Vision Conference 2017, BMVC 2017*, 1–13. <https://doi.org/10.5244/c.31.120> [viewed on 17/04/2020]
- Gupta, U., & Chaudhury, S. (2016). *Deep transfer learning with ontology for image classification*. *2015 5th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, NCVPRIPG 2015*, 1–4. <https://doi.org/10.1109/NCVPRIPG.2015.7490037> [viewed on 17/04/2020]
- Huang, S. (2019). *Building segmentation in oblique aerial imagery*. 58 [viewed on 17/04/2020].
- Kolenbrander, T., Ruiters, F. De, & Yue, T. (2017). *Facade labelling using neural networks*. *Delft University of Technology*. [viewed on 01/02/2021]

- Laupheimer, D., Tutzauer, P., Haala, N., & Spicker, M. (2018). *NEURAL NETWORKS for the CLASSIFICATION of BUILDING USE from STREET-VIEW IMAGERY. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. <https://doi.org/10.5194/isprs-annals-IV-2-177-2018> [viewed on 17/04/2020]
- Laupheimer, Dominik, & Haala, N. (2018). *Deep Learning for the Classification of Building Facades*. 38. *Wissenschaftlich-Technische Jahrestagung Der DGPF Und PFGK18 Tagung in München, 19, 701–709*. [viewed on 17/04/2020]
- Liu, H., Zhang, J., Zhu, J., & Hoi, S. C. H. (2017). *Deepfacade: A deep learning approach to facade parsing. IJCAI International Joint Conference on Artificial Intelligence, 2301–2307*. <https://doi.org/10.24963/ijcai.2017/320> [viewed on 17/04/2020]
- Lotte, R. G., Haala, N., Karpina, M., de Aragão, L. E. O. e. C., & Shimabukuro, Y. E. (2018). *3D façade labeling over complex scenarios: A case study using convolutional neural network and structure-from-motion. Remote Sensing, 10(9)*. <https://doi.org/10.3390/rs10091435> [viewed on 17/04/2020]
- Meltser, R. D., Banerji, S., & Sinha, A. (2018). *What's that Style? A CNN-based Approach for Classification and Retrieval of Building Images. 2017 9th International Conference on Advances in Pattern Recognition, ICAPR 2017, 9–14*. <https://doi.org/10.1109/ICAPR.2017.8593206> [viewed on 17/04/2020]
- Obeso, A. M., Benois-Pineau, J., Acosta, A. Á. R., & Vázquez, M. S. G. (2016). *Architectural style classification of Mexican historical buildings using deep convolutional neural networks and sparse features. Journal of Electronic Imaging, 26(1), 011016*. <https://doi.org/10.1117/1.jei.26.1.011016> [viewed on 17/04/2020]
- Para, W. (2019). *Facade Segmentation in the Wild*. King Abdullah University of Science and Technology Thuwal, Kingdom of Saudi Arabia. [viewed on 17/04/2020]
- Pesto, C. (2016). *Classifying U.S. Houses by Architectural Style Using Convolutional Neural Networks. 1–9*. <http://cs231n.stanford.edu/reports/2017/pdfs/126.pdf> [viewed on 17/04/2020]
- Petrovska, B., Zdravevski, E., Lameski, P., Corizzo, R., Štajduhar, I., & Lerga, J. (2020). *Deep learning for feature extraction in remote sensing: A case-study of aerial scene classification. Sensors (Switzerland), 20(14), 1–22*. <https://doi.org/10.3390/s20143906> [viewed on 01/02/2021]
- Pezzica, C., Schroeter, J., Prizeman, O. E., Jones, C. B., & Rosin, P. L. (2019). *Between images and built form: Automating the recognition of standardised building components using deep learning. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(2/W6), 123–132*. <https://doi.org/10.5194/isprs-annals-IV-2-W6-123-2019> [viewed on 17/04/2020]
- Ramadanta, A. (2010). *KAJIAN TIPOLOGI DALAM PEMBENTUKAN KARAKTER VISUAL DAN STRUKTUR KAWASAN (Studi kasus: Kawasan Ijen, Malang). Jurnal SMARTEK, Vol. 8, No, 130–142*. [viewed on 17/04/2020]
- Sarkar, D., Bali, R., Sharma, T., Sarkar, D., Bali, R., & Sharma, T. (2018). *Deep Learning for Computer Vision. In Practical Machine Learning with*

- Python*. [https://doi.org/10.1007/978-1-4842-3207-1\\_12](https://doi.org/10.1007/978-1-4842-3207-1_12) [viewed on 17/04/2020]
- Sastra, S. (2016). Kajian Estetika Bentuk Pada Fasade Perumahan Real Estate Di Yogyakarta. *Inersia*, 12(1), 78–84. <https://doi.org/10.21831/inersia.v12i1.10355> [viewed on 17/04/2020]
- Taoufiq, S., Nagy, B., & Benedek, C. (2020). *Hierarchynet: Hierarchical CNN-based urban building classification*. *Remote Sensing*, 12(22), 1–20. <https://doi.org/10.3390/rs12223794> [viewed on 01/02/2021]
- Tutzauer, P., & Haala, N. (2017). *Processing of crawled urban imagery for building use classification*. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(1W1), 143–149. <https://doi.org/10.5194/isprs-archives-XLII-1-W1-143-2017> [viewed on 17/04/2020]
- Yoshimura, Y., Cai, B., Wang, Z., & Ratti, C. (2018). *Deep learning architect: Classification for architectural design through the eye of artificial intelligence*. *Lecture Notes in Geoinformation and Cartography*, 249–265. [https://doi.org/10.1007/978-3-030-19424-6\\_14](https://doi.org/10.1007/978-3-030-19424-6_14) [viewed on 17/04/2020]
- Zhuo, X., Monks, M., Esch, T., & Reinartz, P. (2019). *Facade segmentation from oblique UAV imagery*. *2019 Joint Urban Remote Sensing Event, JURSE 2019*. <https://doi.org/10.1109/JURSE.2019.8809024> [viewed on 17/04/2020]