

"Жасанды Интеллект инженері, ғалымы болу" сериясының алтыншы кітабы

Мағжан
ҚАЙРАНБАЙ

6

МАҒЖАН ҚАЙРАНБАЙ

ЖАСАНДЫ ИНТЕЛЕКТ САЛАСЫ БОЙЫНША МАҚАЛА ЖАЗУ ҚҰРАЛЫ

Оқулық

Алматы
2023

УДК 004.8
ББК 32.813
К23

Пікір білдірушілер:
М.Сүлейменов – Қожа Ахмет Яссауи атындағы Халықаралық қазақ - түрік уеиверситеті Ақпараттық технологиялар департаментінің директоры.
Б.Омаров - Әл-Фараби атындағы Қазақ Ұлттық Университеті Ақпараттық Технологиялар факультеті. Ақпараттық жүйелер кафедрасының доценті.

К Мағжан Қайранбаев
"Жасанды интелект саласы бойынша мақала жазу құралы" оқулық/ Алматы "Лантар books" ЖШС, 2023 – 85 бет.

ISBN 978-601-361-497-7

Ғылым әр қашан да алға қарай жылжу мен дамудың негізгі қозғалтқыш құралы болып табылады. Ғылыми тұрғыдан даму әртүрлі эксперименттер мен солардың нәтижелерін конференцияларда баяндау мен тиісті журналдарда жариялау арқылы іске асады. Сол себепті арнайы ғылыми зерттеулер жүргізіп, олардың нәтижелелерін мақала ретінде жазып шығу — өте маңызды іс. Бір ескертетін жайт — мақаланың өзіне тән жазу ерекшеліктері мен рет-тәртібі болады. Қолыңыздағы бұл кітапта сол рет-тәртібі мен өзіндік заңдылықтармен сізді таныстыратын боламыз. Бұл кітап жасанды интеллект саласы бойынша ғылыми- зерттеу жұмыстарын жазуға үйретеді. Бұл кітаптан алған білімдеріңізді ақпараттық технология салаласының өзге де тармақтарында қолданып, мақала жаза алатын боласыз. Бұл кітапта мақаланың келесі бөліктері толығымен түсіндіріледі: "Мақаланың тақырыбы", "Авторлар тізімі", "Аңдатпа", "Кілтті сөздер", "Кіріспе", "Әдеби шолу", "Зерттеу әдістемесі", "Негізгі бөлім", "Нәтиже және оны талдау", "Қорытынды", "Пайдаланылған әдебиеттер". Оқырманға барынша түсінікті болу үшін, қосымша мысалдар келтіріп отырдық. Бұл кітапта негізгі мәселе ретінде "Суреттің эстетикалық әсемдігін тереңдеп үйрету арқылы баға беру" ғылыми-зерттеу жұмысын ұсына отырып түсіндіреміз. Кітап соңында сол мақала үлгі толықтай жазылып көрсетілген. Қолыңыздағы бұл кітаптың басты мақсаты — сіздерге ғылыми тұрғыдан жоғары дәрежелі және халықаралық деңгейдегі мақалалар қалай жазылатынын өз ана тілімізде, яғни қазақ тілінде түсіндіру. Бұл оқу құралын оқып, өз ғылыми зерттеулеріңіз бойынша ағылшынша ғылыми мақала жазып, оларды тиісті журналдарға немесе арнаулы конференцияларға жібере аласыз. Кітап соңында конференциялар мен журналдарды қалай таңдап алу керектігін көрсететін арнайы нұсқаулықтар берілген. Сонымен бірге мақала жазу барысында Google Scholar, LaTeX сияқты құралдармен қалай жұмыс істеу керектігін де түсіндіретін боламыз. Кітаптың соңында постер мен слайдтарды қалай жасау керектігі мен ғылыми этикаға шолу жасаймыз

УДК 004.8
ББК 32.813

ISBN 978-601-361-497-7

© Мағжан Қайранбаев, 2023
© "Лантар books" ЖШС, 2023

Алғы сөз

Ғылым әр қашан да алға қарай жылжу мен дамудың негізгі қозғалтқыш құралы болып табылады. Ғылыми тұрғыдан даму әртүрлі эксперименттер мен солардың нәтижелерін конференцияларда баяндау мен тиісті журналдарда жариялау арқылы іске асады. Сол себепті арнайы ғылыми зерттеулер жүргізіп, олардың нәтижелелерін мақала ретінде жазып шығу — өте маңызды іс. Бір ескертетін жайт — мақаланың өзіне тән жазу ерекшеліктері мен рет-тәртібі болады. Қолыңыздағы бұл кітапта сол рет-тәртібі мен өзіндік заңдылықтармен сізді таныстыратын боламыз. Бұл кітап жасанды интеллект саласы бойынша ғылыми- зерттеу жұмыстарын жазуға үйретеді. Бұл кітаптан алған білімдеріңізді ақпараттық технология салаласының өзге де тармақтарында қолданып, мақала жаза алатын боласыз. Жалпы алғанда, ақпараттық технология саласында ғылыми мақала жазу мынадай бөліктерден тұрады:

1. Мақаланың тақырыбы.
2. Авторлар тізімі.
3. Аңдатпа.
4. Кілтті сөздер.
5. Кіріспе.
6. Әдеби шолу.
7. Зерттеу әдістемесі.
8. Негізгі бөлім.
9. Нәтиже және оны талдау.
10. Қорытынды.
11. Пайдаланылған әдебиеттер.

Мақала әрдайым осы форматта болуы шарт емес. Әйтсе де ғылыми мақалалар көбінесе осы форматта жазылады. Келесі бетте мақаланың жазылу жолы және

оның форматы көрсетілген. Сол формат бойынша біз мақаланың әрбір бөлігін жеке-жеке түсіндіретін боламыз. Сонымен бірге айтпақ ойымыз оқырманға барынша түсінікті болу үшін, қосымша мысалдар келтіріп отырамыз. Бұл кітапта негізгі мәселе ретінде "Суреттің эстетикалық әсемдігін тереңдеп үйрету арқылы баға беру" ғылыми-зерттеу жұмысын ұсына отырып түсіндіреміз. Кітап соңында сол мақала үлгі толықтай жазылып көрсетілген. Ақпараттық технология саласындағы үздік мақалалар ағылшын тілінде жазылады. Сол себепті жоғары дәрежелі ғылыми-мақала жазып шығарғыңыз келсе, онда тек ағылшын тілінде жазу керек болады. Сондықтан біздің ұсынған мысалдарымыз да ағылшын тілінде берілді. Жоғарғы деңгейдегі әрі аса сапалы мақала жазу үшін міндетті түрде ағылшын тілін білу қажет емес, бірақ ағылшын тілінде жазылған мақала оқырмандардың алдына тез жетіп, мейлінше дұрыс бағаланады. Қолыңыздағы бұл кітаптың басты мақсаты — сіздерге ғылыми тұрғыдан жоғары дәрежелі және халықаралық деңгейдегі мақалалар қалай жазылатынын өз ана тілімізде, яғни қазақ тілінде түсіндіру. Бұл оқу құралын оқып, өз ғылыми зерттеулеріңіз бойынша ағылшынша ғылыми мақала жазып, оларды тиісті журналдарға немесе арнаулы конференцияларға жібере аласыз. Кітап соңында конференциялар мен журналдарды қалай таңдап алу керектігін көрсететін арнайы нұсқаулықтар берілген. Сонымен бірге мақала жазу барысында Google Scholar, LaTeX сияқты құралдармен қалай жұмыс істеу керектігін де түсіндіретін боламыз. Кітаптың соңында постер мен слайдтарды қалай жасау керектігі мен ғылыми этикаға шолу жасаймыз.

Жасанды интеллект саласы бойынша ғылыми мақала жазудың құрылымы



1. Мақаланың тақырыбы

Ғылыми жұмыстың тақырыбының аты нақты жазылу керек. Кіріспеде берілген мысал үшін алдымен шешіліп жатқан мәселені айқындап алған дұрыс. Біздің жағдайымызда басты мәселе — фотосуреттің эстетикалық әсемдігіне баға беру, яғни сурет қаншалықты әсем немесе ұсқынсыз екенін анықтау. Шығарып жатқан мәселе айқындалғаннан кейін, ол мәсәленің шешімін табу жолын қарастыру керек. Біздің жағдайымызда "тереңдеп үйретуді және Convolutional Neural Network-ті" пайдала отырып шешу болып табылады. Тақырып толығымен анықталу керек болғандықтан, фотосурет типін сипаттап кеткен жөн. Біздің жағдайымызда "қарапайым фотосуреттермен" емес, тек "портреттермен" ғана жұмыс жасаймыз делік. Осы мәселені шешу барысында қолданылатын амал "compositional augmentation" болсын. Осылайша, тақырыпты толығымен сипаттауға мүмкіндік болып жатса, барынша сол ақпараттарды тақырыпқа қосу керек. Осы мысал негізіндегі мақаланың толыққанды тақырыбы мынадай болады:



Мысал:
Aesthetic Evaluation of Facial Portraits Using Compositional Augmentation for Deep Convolutional Neural Networks

2. Авторлар тізімі

Мақаланы жазуға өз үлесін қосқан барлық адамдар авторлар тізіміне енгізілуі керек. Мақаланы жазып шығуға ешбір үлес қоспаған адам автор ретінде көрсетілмейді. Авторлар тізіміндегі реттілікті сақтау маңызды. Бұл тізім мақалаға қосқан үлесі бойынша дейтін реттілікпен жазылу керек. Мысалы: мақаланың жазылуына ең көп үлек қосқан адам — бірінші, сол үлесі бойынша екінші адам — екінші, тағы сол сияқты болып беріледі. Кейбір жағдайда мақаланың жазылуын ұйымдастырған, бірақ жазылуына қосқан үлесі аз адам мақаланың жетекші авторы ретінде тізімде соңғы болып тұрады. Кей жағдайда осылайша ғылыми-мақаланың жетекшісі кім екенін анықтауға болады.

Аторлардың аты-жөні жазылып болғаннан кейін олардың қай университеттің немесе ғылыми орталықтың өкілі екенін көрсету керек. Яғни, келесі жолға ғылыми бөлімшенің немесе факультеттің атын жазу керек. Одан кейінгі жолға университеттің немесе ғылыми орталықтың атын және мекенжайын жазу керек. Өз өкілеттігін көрсеткен соң әр автордың электронды поштасы жазылады. Көп жағдайда университеттің немесе ғылыми орталықтың электронды поштасын көрсеткен дұрыс.

Негізінде авторлар тізімінің жазылу форматы әрбір журнал немесе конференцияға байланысты әртүрлі болады. Сол себепті өзі қатысып жатқан конференцияның немесе журналдың форматына негізделіп жазылуы керек. Көп жағдайда форматтар бір-біріне ұқсас болып, біз ұсынғандай үлгіде болады.



Мысал:
Magzhan Kairanbay, John See, Lai-Kuan Wong
Center for Visual Computing, Faculty of Computing and Informatics
Multimedia University, 63100 Cyberjaya, Malaysia
magzhan.kairanbay@gmail.com,johnsee@mmu.edu.my,lkwong @mmu.edu.my

3. Аңдатпа

Аңдатпаны көбіне мақала толық жазылып болғаннан кейін жазады. Аңдатпа — мақаланың қысқаша нұсқасы. Мақаланы толық оқу немесе оқымау үшін адамдардың көбі аңдатпаны оқып-ақ шешім қабылдай береді. Сол себепті оның барынша дұрыс жазылуы өте маңызды. Аңдатпаны жазудың құрылымы төмендегідей:

- Шешіліп жатқан мәселені қысқаша сипаттап, оның әсерін атап көрсету кету. Мысалы: *Digital facial portrait photographs make up a massive portion of photos in the web.*
- Шешіліп жатқан мәселені басқалар қалай шешкені туралв қысқаша жазу. Мысалы: *A number of methods for evaluating the aesthetics of photographs have been proposed recently.*
- Өзіңіз ұсынған мәселені шешу жолын қысқаша сипаттап, оның басқа шешімдерден артықшылықтарын көрсету. Мысалы: *However, there have been a little work in the research community to address the aesthetics of targeted image domain, such as portraits. This paper introduces a new compositional-based augmentation scheme for aesthetic evaluation of portraits by well-known deep convolutional neural network (DCNN) models. We present a set of feature augmentation methods that take into account compositional photographic rules to ensure that the aesthetic in portraits are not hindered by standard transformations used for DCNN models.*

- Қол жеткізген нәтижелерді қысқаша сипаттау. Мысалы: *On a portrait subset of the large-scale AVA dataset, the proposed approach demonstrated a reasonable improvement in classification performance over the baseline and vanilla deep learning approaches.*
- Ғылыми жұмыстың қосқан үлесін атап көрсету.



Мысал:

Abstract. *Digital facial portrait photographs make up a massive portion of photos in the web. A number of methods for evaluating the aesthetics of photographs have been proposed recently. However, there have been a little work in the research community to address the aesthetics of targeted image domain, such as portraits. This paper introduces a new compositional-based augmentation scheme for aesthetic evaluation of portraits by well-known deep convolutional neural network (DCNN) models. We present a set of feature augmentation methods that take into account compositional photographic rules to ensure that the aesthetic in portraits are not hindered by standard transformations used for DCNN models. On a portrait subset of the large-scale AVA dataset, the proposed approach demonstrated a reasonable improvement in classification performance over the baseline and vanilla deep learning approaches.*

4. Кілтті сөздер

Кілтті сөздер мақаланың тақырыбын ашатын және ғылыми жұмыстың маңыздылығын көрсететін сөздер болып табылады. Кілтті сөздер немесе сөз тіркестерінің саны тым көп (5-тен көп) немесе тым аз (біреу немесе екеу-ақ) болмағаны дұрыс. Кілтті сөздерді жазу форматы өзі қатысып жатқан конференцияның ерекшелігі немесе журнал форматы бойынша жазылуы керек. Кейбір ғылыми мақалаларда кілтті сөздерді жазу керек болмайды.

Мысал:
Key words: Aesthetics, portraits, deep learning, Convolutional Neural Networks

5. Кіріспе

Кіріспенің басты мақсаты — оқырманды шешіліп жатқан мәселемен таныстыру. Алдымен шешіліп жатқан мәселе қысқаша сипатталып кеткені жөн. Мақаладағы басты проблеманы сипаттау барысында: бұл мәселе қалай туындады, қайдан шықты, қаншалықты жиі кездесетін мәселе және осы мәселе қоғамға қаншалықты ауырпашылық әкелетінін атап көрсеткен дұрыс. Бұларды көрсету барысында статистикалық мәліметтерге немесе өзге де мақалаларға жүгініп, оларға сілтеме беру керек. Жоғарыдағы мысалда бірінші абзацта шешіліп жатқан мәселенің сипаттамасын көре аласыздар: *The ubiquity of smartphones and popularity of social media networks have contributed to an increasing number of photographs, particularly self-portraits or "selfies". A recent study by Bakhshi et al. [1] showed that photos with faces are more popular than other types of pictures. Nevertheless, casual photographers tend to produce many unprofessionally taken photos that are less aesthetically pleasing, and proceed to share them with other people in their circles. As such, an automatic assessment of the photo aesthetics would be useful to provide advice as to which of these photos should be retaken. This is also helpful for designers to select aesthetically pleasing photos for magazine covers, books, and news articles. However, different type of photos needs different type of photographic rules. Even, the most applicable rules, such as rule of thirds [2] could be applied differently for portraits and other type of photos [3].*

Кіріспедегі мәселенің сиппатамасын жазу барысында өз ойыңыздан бұл мәселенің өмірде бар не жоғын болжап жазуға болмайды. Жоғырады айтып өткеніміздей мәселені сипаттағанда басқа да мақалаларға немесе статистикалық мәліметтерге жүгіне отырып, оларды сілтеме түрінде көрсету керек.

Кіріспеде шешіліп жатқан мәселені басқалар қалай шешіп жатқанын сілтемелер арқылы қысқаша сипаттай кеткен жөн. Бұл жағдайда басқалар ұсынған шешу жолдарын біріктіріп, топтастыра сипаттап жазған дұрыс. Басқа шешімдерді сипаттау барысында олардың кемшіліктерін көрсетіп, кейін сіздің ұсынған шешіміңіздің артықшылығын атап көрсету керек.

Шешіліп жатқан мәселеде сізді мотивацияландыратын нәрселер болса, онда оларды сипаттап жазған дұрыс. Мысалы, қандай да бір мәселені шешу барысында басқалар қолданып көрмеген шешу жолын қолданғыңыз келсе, бұл сізді мотивацияландыруы мүмкін. Бірақ, бұл жағдайда үнемі басқа жұмыстарға сілтеме жасап, өзіңіздің мотивацияңыздың барынша маңызды екенін дәлелдеуге тырысу керек. Мысалы: *Deep learning have been widely used in computer vision and natural language processing fields in recent years. In particular, deep convolutional neural network (DCNN) models have made revolutionary steps in object detection and recognition tasks [4–6]. The usage of graphic processing units (GPUs) have made it possible to handle large amounts of data, which directly provides better generalization and discrimination of models. Despite these technological advances and a growing research interest in computational aesthetics ever since the work of Datta et al. [7], very little work has been done in the specific domain of facial or portrait aesthetics, with a number of works treating the aesthetic problem generically [8–10] or categorically depending on its content [11] or query [12].*

Ғылыми диссертация жазу барысында әрбір ғылыми-зерттеу жұмысының мақсат-міндеттеріне мақсатына сәйкес мотивацияландыратын түйінде ойды жазу керек. Әрбір зерттеудің мақсатына сәйкес келетін мотивация мына төмендегідей болуы мүмкін:

Бірінші мотивация: *Most researchers (Lu et al., 2014; Jin et al., 2016) believed that the aesthetic appeal of photographs depends on both low level and high level visual features. Example of low level visual features include edges, contrast, saturation, etc. On the other hand, global high level visual features include placement of objects and relationship between objects as guided by a set of aesthetic photographic rules. Before the revolutionary step of deep learning models, the research community employed handcrafted methods by extracting low level as well as highlevel visual features for classification of aesthetics. However, these attempts may not be sufficiently comprehensive in capturing all lowlevel visual cues and highlevel photographic rules. The innovation of deep learning models, specifically CNNs, allows the extraction of deep features, which can potentially capture a more comprehensive set of features that constitutes both low and high level visual features. In a pioneering attempt, Lu et al. (Lu et al., 2014) proposed the use of two CNNs, for extraction of low level and high level features respectively. The feature extracted from these two CNNs are then merged in the later layers for aesthetics prediction. Jin et al. (Jin et al., 2016) shared the same idea and constructed a CNNbased architecture which uses low level and highlevel visual features for evaluating the photo aesthetics. However, in comparison to Lu et al. (Lu et al., 2014), Jin et al. (Jin et al., 2016) used only one CNN and merged the low level visual features from the earlier layers to the highlevel visual features from the later layers.*

Екінші мотивация: *On the other hand, some researchers believed that photo aesthetics cannot be considered alone without taking into account the photo semantics and photo style. Neumann et al. (Neumann, Sbert, Gooch, Purgathofer, et al., 2005) and Murray et al. (Murray et al., 2012) believed*

that the aesthetics of photographs is closely correlated with the style of photographs. Joshi et al. (Joshi et al., 2011) stated that the aesthetic level of photographs should be evaluated by taking the the photo subject into consideration. Kao et al. (Kao, He, & Huang, 2017) pointed out that photos from different categories should follow different aesthetic criteria. In the same manner, photo aesthetics could be influential at predicting the photo style, semantics and attributes.

Үшінші мотивация: *Very little research works (Broer et al., 2014; Masuda, Gonzalez, Kwan, & Nisbett, 2008; Cela Conde et al., 2009a) consider linking photographic aesthetic perception to the demographic profile of photographers and that photo aesthetics could be evaluated differently among different communities. According to Broer et al. (Broer et al., 2014), the notion of beauty is not universal and demographic, geographic, occupational and ethnic dynamics should be taken into consideration for the evaluation of beauty. According to the studies by Masuda et al (Masuda et al., 2008), there is a strong preference divergence for portrait aesthetics between the Western and Asian societies. CelaConde (CelaConde et al., 2009a) discovered that males and females pay attention to different cues while evaluating photo aesthetics.*

Докторлық, магистратуралық немесе бакалавр дәрежесіндегі диссертацияны жазу барысында "зерттеу сұрақтарын" және "зерттеу мақсатын" жазу керек. Зерттеу сұрақтары ғылыми жұмысты жазу барысында қандай сұрақтарға жауап ізделіп жатқанын атап көрсетеді. Зерттеу сұрақтары көбінше ғылыми мотивацияны көрсетуден кейін пайда болады. Мысалы: *Motivated by the aforementioned problems in the research area at large, we intend to explore a few research questions that warrants further investigation. They can be loosely constructed as follows:*

- *How can we leverage both low level and high level visual features in an efficient manner to classify the aesthetic appeal of photographs?*

- *How compositional rules be applied to show the overall aesthetics without losing the fine grained low level visual feature?*
- *Is the information from the aesthetic evaluation task beneficial to the classification of style and semantics of photographs? Can all these tasks be solved simultaneously rather than the conventional way of considering them separately?*
- *Can image aesthetics be analysed and examined with consideration of the photographer's background? Do the demographic attributes of a photographer influence the aesthetic appeal and style of photographs? Can these demographic attributes contribute towards better prediction of the aesthetic appeal and style of photographs?*

Докторлық дәрежеде диссертация жазу барысында кем дегенде 3-4 ғылыми сұрақ болғаны дұрыс. Ал магистратура және бакалавр дәрежесіндегілерге 1-2 ғылыми сұрақ жеткілікті болады.

Ғылыми сұрақтар қойылғаннан кейін, сол ғылыми сұрақтарға жауап іздеу керек. Ол жауаптар ғылыми жұмыстың алдына қойған мақсаты болып табылады. Ғылыми жұмыстың мақсаттары толығымен анықталуы және дәлме-дәл сипатталуы керек. Ғылыми жұмыстың мақсаттары — шешіліп жатқан мәселе, оның шешілу жолы және сол шешілу жолына қатысты басқа да параметрлермен сипатталады. Жоғардағы ғылыми сұрақтарды шешу үшін алға мына төмендегідей ғылыми мақсаттар қойылуы мүмкін. Мысалы:

- *To design a deep neural network model (шешу жолы) which assesses the aesthetic appeal of photographs (шығарылып жатқан мәселе) by taking into consideration both low level and high level visual (шешу жолындағы параметрлер) features in an efficient manner.*
- *To show that aesthetic features (шешу жолындағы параметрлер) could aid the prediction of photo styles, semantics and attributes (жетілдіру), as well as to facilitate a feasible multi task deep neural network model (шешу жолы).*

- *To investigate the influence of photographers' demographic attributes by the establishment of a new auxiliary dataset (шешу жолы) to better evaluate its impact on the performance of image aesthetic evaluation (шығарылып жатқан мәселе).*

Ғылыми-зерттеудің басты және қосалқы мақсаттары болуы мүмкін. Жоғардағы мысалдағылар басты мақсаттар болса, қосалқы мақсаттар мына төмендегідей болуы мүмкін:
In this thesis, a number of secondary objectives were also explored, particularly: 1) how compositional photographic rules can be shaped into an augmentation strategy to further improve the aesthetic evaluation in specific types of photos such as portraits, and 2) how a variety of machine learning tasks can be exploited by the creation of the new AVAPD dataset. Оқырман әдетте аңдатпаға, кіріспеге және ғылыми-зерттеудің мақсатына баса мән береді. Сол себепті, осы бөлімдердің дұрыс жазылуына баса көңіл бөлу керек.

Шешіліп жатқан мәселені, өз мотивацияңызды және басқа да шешілу жолдарын көрсеткеннен кейін барып өзіңіздің ұсынып жатқан шешіміңізді қысқаша сипаттап кеткеніңіз дұрыс. Сіздің ұсынған шешіміңіздің басқалар ұсынған шешімдерден артықшылығын сипаттап жазыңыз. Сонымен бірге ұсынып отырған шешіміңіздің сипаттамасын визуалды тұрғыда көрсетіп кеткеніңіз жөн болар. Өзіміз ұсынған мысал бойынша шығарылған шешімінің сипаттамасы және соған сәйкес оның визуалды кескіні былайша берілген (1-сурет): *Motivated by these developments, we are inspired to improve existing deep learning techniques to better evaluate facial aesthetics. In this paper, we present a new compositional-based augmentation scheme for deep convolutional neural networks (DCNN) that can boost the aesthetic evaluation of facial portraits. Conventional sampling of sub-images (i.e. warping, center-crop, random crop) as receptive fields of a DCNN is rigid and can potentially distort the composition in portraits (Figure 1).*

Темірбек Жүргенов атындағы
Қазақ ұлттық өнер академиясы
ИНВ № 104927
Казахская национальная академия искусств
имени Темирбека Жургенова

As such, we exploit the rule of thirds composition rule to generate a feasible set of inputs to the network, which are in turn aggregated to form a deeply-learned augmented feature set. On a portrait-based dataset derived from the large-scale AVA dataset, our approach is able to surpass the performance of handcrafted and compositional-based features [3, 13] which have been a staple for defining beauty in portraits. The objective of this paper is to provide the methods of evaluating the portraits aesthetics using deep neural networks.

Өзіңіздің ұсынған шешіміңізді көрсеткеннен кейін осы ғылыми жұмыс арқылы қосқан өз үлесіңізді жазып қойсаңыз болады. Әрбір зерттеу мақсатына сәйкес қосқан үлесіңді атап көрсету керек. Біздің зерттеу жұмысымыз да үшін қосқан үлесіміз мына төмендегідей: *There are three main contributions of this thesis. Each of them are elaborated below.*

One of the contributions of this research work is the usage of Global Average Pooling (GAP) layer to decrease the model size as well as produce the Class Activation Maps(CAM). CAM can illustrate the regions of aesthetic appeal in the form of a heat map. The GAP layer is connected to the convolutional layers of CNN to produce feature vectors at different levels. The merging of these feature vectors are used as descriptors for classifying the aesthetic level photographs. Intuitively, the aggregation of features from the earlier layers as well as from the later layers are useful to extract both low and highlevel visual features from the photographs. The proposed approach shows the comparatively same results as baseline solutions but uses less amount of parameters.

To the best of our knowledge, there is no deep learning model which can solve three different tasks such as photo aesthetic evaluation, photo style and semantic prediction in one network. This research work proposes the joint deep learning model which can solve the three independent tasks mentioned above in one network with better accuracy at predicting photo semantics, styles and attributes.

Another contribution of this research work is the creation of the AVAPD dataset which is the extended version of AVA dataset (Murray et al., 2012). The AVAPD dataset contains the demographic attributes of photographers based on the images in AVA dataset. The combination of AVA images with metadata from AVAPD was able to outperform the accuracy of baseline solutions for the photo aesthetic classification task. In the same manner, the performance of style classification task is also improved, because of additional demographic metadata.

Кіріспенің соңында келесі параграфтарда оқырманды не күтіп тұрғанын қысқаша жазып қоюға болады. Мысалы: *The following sections are organized as follows: Section 2 reviews the current works in literature, Section 3 describes the dataset that was used, Section 4 elaborates on the proposed method in detail, Section 5 reports the experimental results with further discussion on them, while Section 6 concludes the paper.*

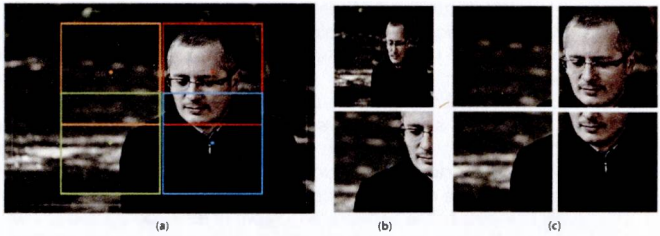


Fig. 1. Compositional-based cropping for DCNN inputs: (a) A portrait can be cropped aesthetically from power points defined by rule of thirds composition principle, (b) Full image warping (top) and center-cropping (bottom) can typically distort composition in portraits. (c) The four sub-images based on the power points.

<u>Толық кіріспеге арналған мысал:</u>
The ubiquity of smartphones and popularity of social media networks have contributed to an increasing number of photographs, particularly self-portraits or "selfies".



A recent study by Bakhshi et al. [1] showed that photos with faces are more popular than other types of pictures. Nevertheless, casual photographers tend to produce many unprofessionally taken photos that are less aesthetically pleasing, and proceed to share them with other people in their circles. As such, an automatic assessment of the photo aesthetics would be useful to provide advice as to which of these photos should be retaken.

This is also helpful for designers to select aesthetically pleasing photos for magazine covers, books, and news articles. However, different type of photos needs different type of photographic rules. Even, the most applicable rules, such as rule of thirds [2] could be applied differently for portraits and other type of photos [3].

Deep learning have been widely used in computer vision and natural language processing fields in recent years. In particular, deep convolutional neural network (DCNN) models have made revolutionary steps in object detection and recognition tasks [4–6]. The usage of graphic processing units (GPUs) have made it possible to handle large amounts of data, which directly provides better generalization and discrimination of models. Despite these technological advances and a growing research interest in computational aesthetics ever since the work of Datta et al. [7], very little work has been done in the specific domain of facial or portrait aesthetics, with a number of works treating the aesthetic problem generically [8–10] or categorically depending on its content [11] or query [12].

Motivated by these developments, we are inspired to improve existing deep learning techniques to better evaluate facial aesthetics. In this paper, we present a new compositional-based augmentation scheme for deep convolutional neural networks (DCNN) that can boost the aesthetic evaluation of facial portraits. Conventional sampling of sub-images (i.e. warping, center-crop, random crop) as receptive fields of a DCNN is rigid and can potentially distort the composition in portraits (Figure 1). As such, we exploit the rule of thirds composition rule to generate a feasible set of inputs to the network, which are in turn aggregated to form a deeply-learned augmented feature set. On a portrait-based dataset derived from the large-scale AVA dataset, our approach is able to surpass the performance of handcrafted and compositional-based features [3, 13] which have been a staple for defining beauty in portraits. The objective of this paper is to provide the methods of evaluating the portraits aesthetics using deep neural networks.

The following sections are organized as follows: Section 2 reviews the current works in literature, Section 3 describes the dataset that was used, Section 4 elaborates on the proposed method in detail, Section 5 reports the experimental results with further discussion on them, while Section 6 concludes the paper.

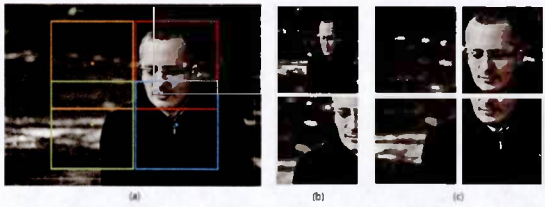


Fig. 1. Compositional-based cropping for DCNN inputs: (a) A portrait can be cropped aesthetically from power points defined by rule of thirds composition principle, (b) Full image warping (top) and center-cropping (bottom) can typically distort composition in portraits. (c) The four sub-images based on the power points.

6. Әдеби шолу

Кіріспеде өзге де шешімдерге қысқаша шолу жасап, олардың кемшіліктерін атап өту арқылы өзіміздің мәселені шешу жолымызды таныстырған болатынбыз. Осыдан кейін толықтай әдеби шолуға кірісуімізге болады. Әдеби шолуда басқа шешімдердің бәрін тізбектей атап өтпей-ақ, оларды өзара топтастыру арқылы таныстырған жөн. Топтастыру хронологиялық реттілікпен немесе мәселені шешу жолына қарай жүргізілуі мүмкін. Мәселені шешу жолына қарай дейтін ретпен топтастыруда алдымен мәселені бірінші шешу жолына қатысты шешімдерді жазып шығып, кейін екінші шешу жолына қатысты шешімдерді таныстыру, тағы сол сияқты жолмен топтастырылады. Мысалы: 1) *Works from the early efforts of Datta et al. [7] to the recent use of popular descriptors [8] have concentrated on the extraction of hand-crafted visual features for aesthetic evaluation of photos. 2) Different forms of convolutional neural networks [15, 10, 16] have been applied to solve this task, with the intent of learning features in a more natural and unguided manner.* Әрбір шешу жолына қатысты қолданылған әдебиеттерді таныстырғанда оларды сипаттап қана қоймай, артықшылықтары мен кемшіліктерін де атап өткен дұрыс. Мысалы: *The proposed method (after shrinking to a selected set of 12 positively-correlated features) outperformed the baseline method of [3] by 14.66% on the same data. However, using only face-related features without any selection, it performed better by only 1%.*

Another work by Lienhart et al. [17] applied similar types of features on three small datasets, while also performing aesthetic score prediction by support vector regression. But their work does not have the similar rigor and extensiveness as that of [13]. Әдеби шолуды осындай реттілікпен жасап шыққаннан кейін өзіңіз ұсынған шешімді қысқаша (бірнеше сөйлеммен ғана) сипаттап, өз шешіміңіздің басқалардан айырмашылығын (мүмкін артықшылығын) атап көрсеткен дұрыс.



Мысал:

The task of evaluating the aesthetic quality of photographs is becoming increasingly popular in recent years. As the notion of aesthetics is a subjective one, substantial human effort to rate photos is needed. Nonetheless, the creation of various datasets such as CUHKPQ [11] and AVA [9], from online photo communities such as Photo.net and DPChallenge [14], have made it possible to evaluate solutions with ground truth labels. Works from the early efforts of Datta et al. [7] to the recent use of popular descriptors [8] have concentrated on the extraction of hand-crafted visual features for aesthetic evaluation of photos. In [7], as many as 66 pre-determined features, depicting colourfulness, composition, textures and statistical properties were crafted while [8] went for generic representations such as SIFT and Color descriptors, supplemented by Bag-of-Words (BoW) and Fisher Vector (FV) encodings and spatial pyramids (SP).

Different forms of convolutional neural networks [15, 10, 16] have been applied to solve this task, with the intent of learning features in a more natural and unguided manner. We take particular interest in the work of Lu et al. [10], which generated two heterogeneous inputs to represent global and local cues respectively. Global views are represented by normalized inputs such as center-crop, warp and padding, while local views are represented by randomly-cropped inputs of smaller parts of the image. These views were used generically for any type of photo, and may not be suitable for photos with facial portraits. However, a majority of DCNN-based works showed scant attempts to consider compositional impact relating to photographic rules.



Мысал:

However, not many works focused on the aesthetics of photos from a specific domain, particularly in facial portraits. The works in [3, 13, 17] try to rate the aesthetic quality of face portraits and understand the visual features that define the beauty of digital portraits, while another recent work [18] attempted to answer the question of "how to take a good selfie?" from the perspective of image popularity.

Khan and Vogel [3] made a comparative study on the use of traditional features [7] versus their choice of features that are focused on spatial and color composition. Firstly, template based spatial composition features were constructed from three templates for power points, power lines and their combination which were modified from [19]. Secondly, highlight and shadow composition features were obtained by extracting the face illumination, background contrast, brightness and size of the face as features. Their results show that the proposed features are more informative, more concise and better equipped to rank portraiture composition. The proposed features are more accurate by 6% and smaller 9 times in feature length compared to traditional features. More importantly, the authors postulated that spatial composition is more important than highlight and shadow composition when identifying aesthetics.

Redi et al. [13] introduced their own features to describe the portrait aesthetics. These features have been derived from the following photographic dimensions: compositional rules (lighting, sharpness, traditional compositional features), semantics and scene content (object bank feature), basic quality metrics (noise, exposure quality, JPEG quality), portrait specific features (facial landmark sharpness, landmark statistics and face/background contrasts) and fuzzy properties (emotion, originality, memorability and uniqueness). Their feature analysis showed that race, gender and age of a person are uncorrelated with photographic beauty, while soft and hard image properties such as sharpness of facial landmarks, image contrast, exposure, uniqueness and originality are directly related to portrait aesthetics. The proposed method (after shrinking to a selected set of 12 positively-correlated features) outperformed the baseline method of [3] by 14.66% on the same data. However, using only face-related features without any selection, it performed better by only ~1%. Another work by Lienhart et al. [17] applied similar types of features on three small datasets, while also performing aesthetic score prediction by support vector regression. But their work does not have the similar rigor and extensiveness as that of [13].

Әдеби шолуды хронологиялық реттілікпен жазу барысында оны визуалды түрде де көрсетсе болады. Мысалы, төмендегі суретте оның әдемілігін (көріктілігін) қарастырып, зерттеген мақалар көрсетілген (2-сурет).

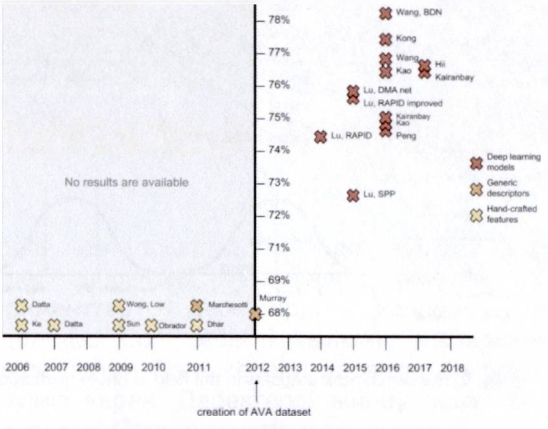


Fig. 2. History of research works which investigate the evaluation of aesthetics of photographs

Кей жағдайда қолданып жатқан деректер қорына шолу жасап, оларды өзара салыстыру керек болады. Салыстыру барысында олады визуалды түрғыда немесе кесте түрінде көрсетуге болады. Мысалы, төмендегі кесте (2.1-кесте) мен суретте (3-сурет) сол деректер қорын өзара салыстырған болатынбыз.

Table 2.1: Comparison of datasets proposed for photo aesthetic evaluation.

Datasets	Number of images	Evaluation score range
PhotoNet	3 581 -	1-7
CUHK	60,000	0 either 1
CUHKPQ	-	-
JenAesthetics	1628	-
MIRFLICKR	-	-
AADB	10,000	0-1
AVA	2500,000	1-10

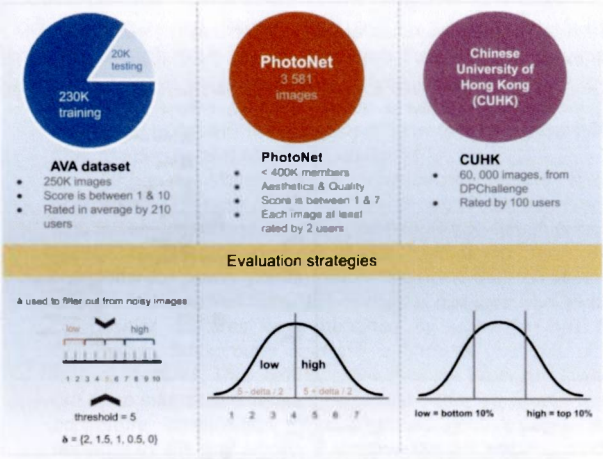


Fig. 3. The benchmark datasets in the field of photo aesthetics

Бұл жағдайда деректер қорының аттары, барлық деректер саны, шындық айнымалыларының мәндер аралықтары, деректерге баға беру жолдары және барлық баға берген адамдардың саны көрсетілген.

7. Зерттеу әдістемесі

Жасанды интеллект саласында зерттеу әдістемесі көбінше деректер негізінде жүреді. Барлық эксперименттер деректерді колдану арқылы орындалады. Сол себепті ғылыми эксперименттерді бастамас бұрын бізге керекті деректердің барына көз жеткізуіміз керек. Деректерді жинау жолы сан түрлі болуы мүмкін. Олардың негізгілері мынадай:

- Сауалнама негізінде жиналған деректер. Адамдарға сауалнама жүргізу арқылы сұрақтар қойып, сұрақ жауаптарын зерттеу негізінде ғылыми жұмыс жасау.
- Ашық деректермен жұмыс істеу. Мұнда басқа ғылыми жұмыстардың қортындысы ретіндегі деректер ашық түрде желіге салынып қойылады. Сол деректерді қолдана отырып ғылыми жұмыс жасауға болады.
- Немесе деректерді өзіңіз жинап, сол деректермен ғылыми-зерттеу жұмысын жүргізуіңізге болады.

Деректерді өзіңіз жинайтын болсаңыз, кей жағдайда сол деректерді шындық мән беруге тура келеді. Мұндай жағдайда өзіңіз сол мәндерді берсеңіз болады немесе достарыңыз сол шындық мәнді сізге беру арқылы көмекке келуі мүмкін. Бірақ, көп ретте көпшіліктің көмегі арқылы amazon mechanical turk сияқты құралдардың көмегімен сол шындық мәндер жинап алынады.

Ашық деректерді қолдансаңыз да немесе өз дерегіңізді қолдансаңыз да, сол деректерді барынша сипаттап жазуға тырысыңыз. Деректерді сипаттау барысында визуалды тұрғыдан сипаттап кетуге де болады (4, 5-суреттер).

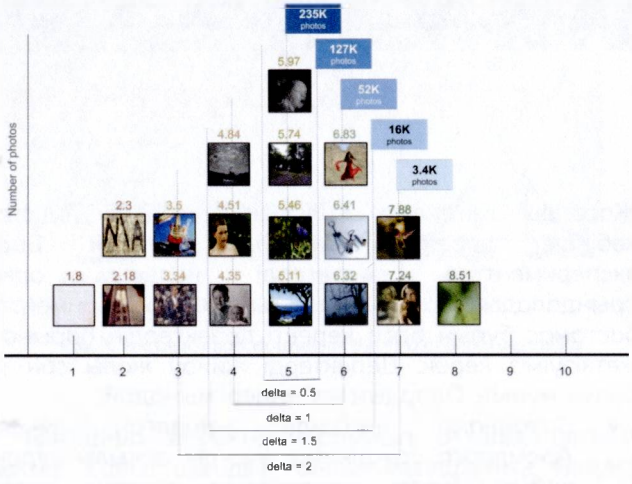


Fig. 4. Samples from AVA dataset for different value of δ

66 Semantic Tags of AVA dataset

Abstract	Cityscape	Horror	Nude	Seascapes	Travel
Action	Digital Art	Humorous	Overlays	Self Portrait	Urban
Advertisement	Diptych / Triptych	Infrared	Panoramic	Sky	Vintage
Analog	DPChallenge GTGs	Insects	Performance	Snapshot	Water
Architecture	Emotive	Interior	Persuasive	Sports	Wedding
Astrophotography	Family	Landscape	Photo-Impressionism	Still Life	
Birds	Fashion	Lensbaby	Photojournalism	Street	
Black and White	Fish Eye	Macro	Pinhole/Zone Plate	Studio	
Blur	Floral	Maternity	Political	Texture Library	
Camera Phones	Food and Drink	Military	Portraiture	Textures	
Candid	HDR	Music	Rural	Traditional Arts	
Children	History	Nature	Science and Technology	Transportation	

Fig. 5. 66 Semantic tags of AVA dataset

Машинаны үйрету салаларында деректерді үйрету үшін оларды валидация және тест жиынтықтарына бөлу керек. Бұл жағдайда деректерді бөлу деректер жиынтығының өлшеміне байланысты болады. Егер сіз 1 000 000 көлеміндегі деректермен жұмыс істесеңіз, онда деректердің 98%-ын үйретуге және 1%-ын валидация және басқа 1%-ын тестілеуге пайдалана аласыз. Ал егер деректер жиынтығы аздау болса (мысалы, 10 000 көлемінде), онда 80%-ын үйретуге, және 10%-ын валидацияға және қалған 10%-ын тестілеуге пайдалана аласыз. Ал егер сіз деректерді басқа мақаладан немесе ресурстан алып қолданатын болсаңыз, онда сол деректердегі үйрету, тестілеу бөліктерін қолдануыңыз керек. Тіпті үйретуге және тестілеуге қолданылған деректерді алмастырмастан, сол реттілікпен қолдану керек. Сол сәтте, сіз ұсынған шешу жолын тестілеу деректеріне пайдаланып, кейін оны басқа шешімдермен салыстыра отырып өз шешіміңіздің артықшылықтарын атап көрсете аласыз. Мысалы, өзіміздің ғылыми зерттеу жұмысында AVA атты деректерді қолдандық. Мұнда біз осы деректер жиынтығын басқа авторлар ұсынған ретпен үйрету және тестілеу бөлігіне бөлдік. Тіпті, авторлармен хабарласып, қай суреттер үйретуге және қай суреттер тестілеуге қолданылатынын анықтап алдық. Осылайша, өзіміздің шешу жолымызды басқалармен әділ түрде салыстыра аламыз.

Деректерді жинап, сипаттап болғаннан кейін олардың кейбіреуін визуалды түрде көрсеткен абзал. Деректерді көрсету барысында шындық мәндерді де көрсеткен дұрыс. Біздің ғылыми жұмысымыздағы AVA және AADB деректер қорындағы әрбір деректер төмендегі суретте көрсетілген (6, 7, 8-суреттер).

Деректер дайын болғаннан кейін оларға эксперимент жасап, оның нәтижесіне қол жеткізу үшін бағалау өлшемдерін қолдану арқылы экспериментке баға беру керек. Бағалау өлшемдері шығарылып жатқан есепке байланысты. Мысалы, классификация есебі үшін дәлдік, нақтылық сияқты өлшем бағалары қолданылады (1-теңдеу). Ал егер регрессия сияқты

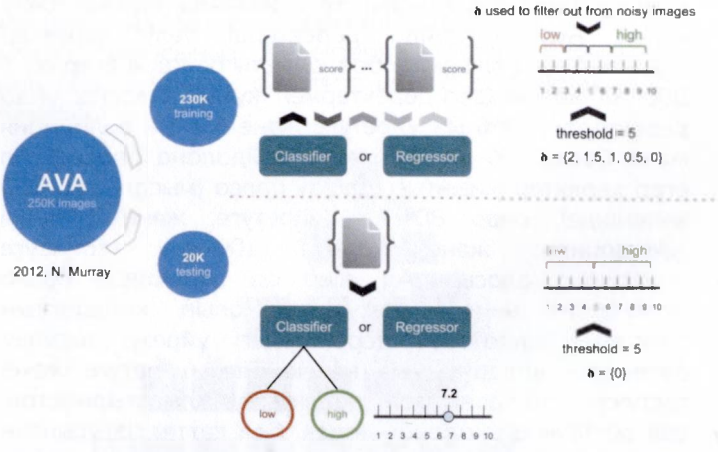


Fig. 6. Training and testing framework in AVA dataset



Fig. 7. Style categories and samples from AVA dataset

$$accuracy = \frac{count(class_{predicted} = class_{actual})}{N} * 100\%$$

$$MeanAbsoluteError(MAE) = \frac{1}{N}\sum |y_i - \hat{y}_i| \quad \rho = 1 - \frac{6\sum d_i^2}{n^3 - n}$$

Eq. 1. Evaluation metrices: top - accuracy, bottom left - Mean Absolute Error, bottom right - Spearman correlation



Fig. 8. The aesthetically low and aesthetically high photos from AADB and AVA dataset

есепті шығарып жатсақ, онда Mean Absolute Error, Mean Square Error сияқты бағалау өлшемдерін қолдана аламыз. Егер басқа мақалалардан алынған деректер қорын пайдалансақ, онда сол мақалаларда пайдаланылған бағалау өлшемдерін қолдану керек. Себебі, өзіміздің ұсынған шешім мен басқа мақалаларда ұсынылған шешімдерді бір деректер қорында және бір бағалау өлшемінде бағалап, салыстыруымыз керек. Жазып жатқан мақаламыздың басты мақсаты — сол деректер қорында және сол бағалау өлшемінде басқалар ұсынған шешімлерден асып түсетін артықша шешім ұсыну.

Мақала жазу барысында бірнеше деректер қорын қолданып, әрбір деректер қоры бойынша басқаларға қарағанда жақсы жұмыс істейтін шешім тауып ұсынуға тырысу керек. Яғни тек бір ғана деректер қорымен жұмыс істеп қоймай, бірнеше деректер қорымен жұмыс істеуге тырысыңыз.

Сонымен бірге әрбір деректер қоры үшін соған сәйкес бағалау өлшемін қолданып, басқа мақалаларда ұсынылған шешімдерді асып түсуге тырысу керек.



Мысал:

The Aesthetic Visual Analysis (AVA) [9] dataset is the largest known dataset that is widely used to evaluate the aesthetic value of the images. The dataset contains more than 250,000 images from DPChallenge [14] with ground truth data such as human ratings ranging from 1 to 10 (with average 210 ratings per photo), with further meta-annotations from 66 textual semantics and 14 photographic styles from 72 different photo challenges. Following the work in [13], we select from the entire AVA dataset photos that are tagged with "Portraits" and also those whose challenge title contained the words Portrait, Portraiture or Portraits. Then, we filter further by detecting for photos that contained the presence of one or more faces with Face++ API [20]. A total of 10,141 images were obtained to form the portrait subset dataset. For better clarity, we refer to this subset as "AVA-Portraits" from this point onwards.

The same evaluation methodology as [13] was used. The dataset is randomly divided into five equal partitions. To ensure fair experimentation, we obtained the partition information from the authors. One partition will be used as test set and the rest as training set. Using 5-fold cross validation, the final accuracy is determined by averaging the minimum and maximum result of these five different cases. A portrait image is classified as aesthetically high if its aesthetic score is more than 5.5 (mean user score) [13]. Otherwise, the image is classified as aesthetically low.

8. Негізгі бөлім

- Негізгі бөлімде өзіңіздің ұсынған шешіміңізді толығымен сипаттап көрсету керек. Егер сізде бірнеше шешім болса, онда басты бөлімді бірнеше бөлікке бөлу арқылы түсіндірген дұрыс болады.
- Негізгі бөлімдегі ұсынған шешімді түсіндіру барысында визуалды диаграммаларды қолданған дұрыс.
- Кей жағдайда математикалық формулаларды немесе псевдо кодты қолдану арқылы түсіндіре аласыз.
- Өзіңіз ұсынған шешімнің аты тым ұзын болған жағдайда оның атын қысқартып жазған дұрыс. Себебі, өз шешіміңізді әрбір атаған сайын оның толық, яғни ұзақ атын жазбай, қысқарған атын жазу ыңғайлы болады.
- Негізгі бөлім мақаланың басым бөлігін қамтуы керек.



Мысал:

Convolutional neural networks (CNN) are biologically-inspired variants of the multilayer perceptron, dating back to the work of LeCun et al. [21]. However, it started to be increasingly popular only in 2012, after the work of Krizhevsky et al.

[4] introduced an eight-layer deep CNN network, winning the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. From then on, convolutional neural network models have shown unrivaled performance in ILSVRC each year. In ILSVRC, models are trained and tested on the ImageNet dataset [22], which consists of more than 14 million of images with 1,000 classes. The classes comprise of a large variety of objects such as car, motocycle, person, cats, etc. Realistically, an immense mount of computational power is necessary to be able to train weights in these deep networks at a reasonable amount of time. Thus, pre-trained models are commonly available for application to different related classification problems. However, in view of the different number of classes in other problems, the number of nodes in the last softmax layer should be replaced with that equivalent to the number of classes. In our case, the softmax layers are 2x1 as it is a binary classification problem (good aesthetics, bad aesthetics) (See Figure 9).

There are a number of popular pre-trained models available for public use. In this work, we use the Overfeat [5] (winner of the localization task in ILSVRC 2013) and GoogleNet [6] (winner of the classification task in ILSVRC 2014). For each of these models, the number of nodes in last (softmax) layer is reduced from 1000 to 2, and the weights are fine-tuned with the new training data. The second last layer of the network is then taken as the feature vector for classification by a softmax classifier (or multi-class binary logistic regression). We use the Sklearntheano [23] library, an open source library on the Theano framework for the implementation of this research.

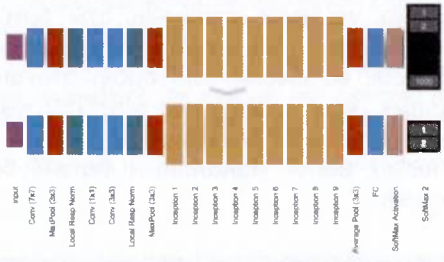


Fig. 9. *Changing the number of nodes in the last layer of GoogleNet.*

4.1 Feature augmentation using photographic rules

Feature augmentation is a process intended to increase the feature vector length at a certain layer by aggregating features obtained from various transformed versions of the input image. We distinguished this from data augmentation, which is a well-known technique [24, 4] used especially for learning deep representations

by generating additional number of training samples. By augmentation we mean perturbing an image by transformations (e.g. cropping, flipping, or a combination), leaving the underlying class unchanged.

In our work, we propose a feature augmentation scheme based on compositional photographic rules. This is motivated by the fact that typical image transformations such as cropping and flipping do not make aesthetic sense when it comes to face portraits. They may aggravate the actual composition of the portrait since a majority of cropping are done randomly or without applying compositional principles, while warping can potentially distort the face aspect ratio. The rule of thirds [19, 2, 25] is one of the popular photographic rules which divides the image to nine parts by 3 x 3 (see Figure 10). The four intersections of the rule of thirds lines are called power points. According to this popular rule, a photo is considered aesthetically appealing if the object of interest is located near one of these four power points.

*We design a number of compositional cropping arrangements that are motivated by photographic rules associated to rule of thirds. Firstly, the baseline transformation involves warping the image to a constant width and height (corresponding to the model) and passing it as a single input to the DCNN model. We denote this in our experimental reporting as '**warp**'.*

*According to [3], the face, which is the main subject of interest in the portrait, should be located at the top power line by keeping the eyes in the upper third of the image area. Meanwhile, the symmetrical composition rule states that it is aesthetically pleasing to center the face in the image frame. Inspired by these photographic composition rules, we modify the rule of thirds for portraits by taking two crops centered on the top two power points (left and right) and a center crop between these two crops (see Figure 10a). All three crops are passed to the DCNN model as individual inputs to generate feature vectors that are then concatenated. We denote this technique as '**c3**'. The third technique (denoted as '**c4**') uses four different sub-images of constant height and width, center cropped from four power points based on the rule of thirds (see Figure 10b). The augmented feature vector is formed by concatenating the four feature vectors generated from each crop.*



Fig. 10. *a) Three-crop by rule of thirds, b) Four-crop by rule of thirds.*

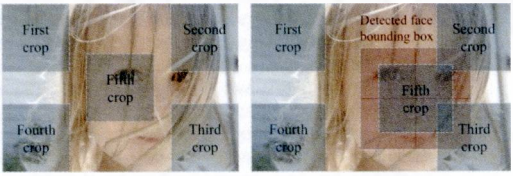


Fig. 11. *a) Five-crop: four crops from four different corners + center crop) b) Five-crop: four crops from four corners + face crop.*

*Extending on to five-crop method, we propose two different variants. The first (denoted as '**c5**') takes four sub-images from the four corners of the image, and the center crop as the fifth. Since the fifth crop provides information from the middle of the photo, the use of four crops from the corners yields better coverage than using crops centered at all four power points. The second variant (denoted as '**c5f**') crops the face region as the fifth crop. Face++ [20] API was used to detect the location of the center coordinate, and a bounding box that defines the sub-image is cropped and squashed to a constant width and height. Figure 12 shows some sample face regions detected by Face++ API and used as the fifth crop for method '**c5f**'. Similar to previous schemes, DCNN features resulting from all five crops are concatenated. Due to slight unreliability of the Face++ API, only about 77.24% of all faces from the AVA-Portraits dataset can be detected. As such, the affected images fall back on the center crop as the fifth crop. Alternatively, we also manually annotated the faces in the remaining affected images. We denote this as '**c5f-m**'.*



Fig. 12. *Sample cropped faces detected by Face++ API*

For all compositional cropping methods, we crop the sub-images based on a fixed width and height of 256 pixels, which are subsequently resized to receptive field sizes of their respective networks i.e. 234 for Overfeat, 224 for GoogleNet, both in RGB color space with zero mean.



Table 2. Aesthetic classification accuracy and dispersion error of the proposed methods compared against baseline methods.

Methods	Accuracy (%)	Error ± (%)
Baseline A [3]	54.585	1.25
Overfeat-warp	63.211	1.48
GoogleNet-warp	63.751	0.85
Overfeat-c3	63.861	0.44
Overfeat - c4	64.133	0.62
Baseline B [13]	64.240	1.76
GoogleNet-c3	64.244	0.64
GoogleNet-c4	64.519	1.06
Overfeat-c5	64.536	1.11
GoogleNet-c5	64.962	1.20
GoogleNet-c5f	**65.216**	1.20
GoogleNet-c5f-m	**65.330**	1.01

9. Нәтиже және оны талдау

Бұл бөлімде эксперимент нәтижесін көрсетіп жатудың қажеті жоқ. Нәтижені көрсетуден басқа, жеткен нәтижеге талдау жасап, өзіңіз ұсынған шешімнің артықшылықтары мен кемшіліктерін көрсету керексіз. Ұсынған шешіміңізді басқа шешімдермен бірге кесте форматында салыстырған дұрыс. Бұл жағдайда салыстыру әділ болуы үшін барлық шешімдерге бір ғана бағалау өлшемі мен дерек қорын қолданған керек (3-кесте).

Table 3. Comparison of proposed solution with other state-of-the-art solutions

Methods	Accuracy (%)	Error ± (%)
Baseline A (Khan & Vogel, 2012a)	54.585	1.25
Overfeat-warp	63.211	1.48
GoogleNet-warp	63.751	0.85
Overfeat-c3	63.861	0.44
Overfeat - c4	64.133	0.62
Baseline B (Redi et al., 2015a)	64.240	1.76
GoogleNet-c3	64.244	0.64
GoogleNet-c4	64.519	1.06
Overfeat-c5	64.536	1.11
GoogleNet-c5	64.962	1.20
GoogleNet-c5f	**65.216**	1.20
GoogleNet-c5f-m	**65.330**	1.01

Өзіңіз ұсынған шешімді қара әріптермен ерекшелеп көрсетсеңіз де болады. Мақала жазу барысында басқа-

Table 4. Comparison of proposed solution with other state-of-the-art solutions, where proposed solution highlighted with bold

Method	Work	Acc. (%)
Baseline (SIFT)	(Murray et al., 2012)	67.00
Generic Descriptors*	(Marchesotti et al., 2011)	68.55
Low level visual features*	(Datta et al., 2006)	68.67
High-level perceptual features*	(Ke et al., 2006)	71.16
RDCNN-Style	(Lu et al., 2014)	74.46
A&C CNN	(Kao, Huang, & Maybank, 2016)	74.51
DMA-Net	(Lu, Lin, Shen, et al., 2015)	75.41
RDCNN-Semantic	(Lu, Lin, Jin, et al., 2015)	75.42
ResNet50 + Triplet Loss	(Schwarz, Wieschollek, & Lensch, 2018)	75.83
MTCNN	(Kao, He, & Huang, 2016)	76.15
Multi-Att AnGAP-FeatEns	(Kairanbay, See, Wong, & Hii, 2017)	76.32
MSDLM	(W. Wang et al., 2016)	76.94
Attribute+Content CNN	(Kong et al., 2016)	77.33
MNA-CNN	(Mai et al., 2016)	77.40
Deep Chatterjee's Machine	(Z. Wang et al., 2017)	78.08
MTRLCNN (ResNet)	(Kao et al., 2017)	79.08
Baseline CNN	**this work**	**76.37**
MDgCNN	**this work**	**78.44**
ADgCNN	**this work**	**78.58**

дардың шешімдерінен асып түсетін шешім ұсынған керек. Көрсеткен нәтиже мен жеткен жетістіктеріңізді мүмкіндігінше визуалды тұрғыда да сипаттап көрсеткен дұрыс. Мысалы, біздің ғылыми-зерттеуімізде ұсынған бірнеше шешіміміздің нақтылығын төмендегідей фигура арқылы көрсетуге болады (13-сурет).

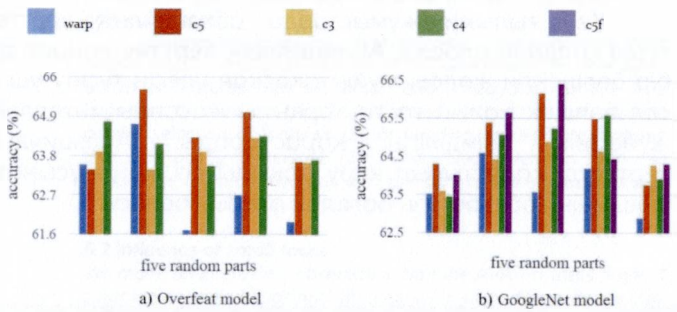


Fig. 13. Accuracy of Overleaf and GoogleNet models with different input formats

Бұл жерде оң және сол жақта екі түрлі шешім шығару моделі берілген. Әрбір шыққан шешім үшін әртүрлі параметрлер ұсынылып, одардың нәтижесі көрсетілген.

Егер сіз суреттермен жұмыс істесеңіз, онда шыққан нәтижелерді көрсете кеткен дұрыс. Біздің ғылыми жұмыста портреттермен жұмыс істеу барысында адамның түрін анықтау керек. Адамның бетін анықтау нәтижелері төмендегі суретте көрсетілген (14-сурет).



Fig. 14. *Detection of faces using Face++ API*

Кей ғылыми жұмыстарда абляциялық зерттеу түрін қолдана аласыз. **Абляциялық зерттеу** қандай да бір бөлшектің жалпы жүйеге қосқан үлесін түсіну үшін, сол бөлшекті алып тастау арқылы жасанды интеллект жүйесінің өнімділігін қарастырады. Абляциялық зерттеуді қарастырып көру арқылы сіз өзіңіз ұсынған шешімнің әрбір бөлігін бағалай алатын боласыз.

Мысли:

5 Results and Discussions

5.1 Overall results

*We define two baselines for the purpose of comparisons– one by [3] which relies on traditional spatial and color composition features (**'Baseline A'**), and another [13] which uses a discriminative set of visual features (**'Baseline B'**). Two DCNN models were utilized in our experiments: the 8-layer **Overfeat** model and the 22-layer very deep **GoogleNet** model.*

Table 1 shows the proposed and baseline approaches on the AVA-Portraits dataset. We achieve the best result reported thus far for this dataset, of 65.33%. From the feature augmentation results, we can see that the aesthetic classification accuracy generally increases as more crops were added to augment the DCNN-learned features (see Figure 16a). The best performing scheme uses five crops, with the variant that takes a face as fifth crop marginally higher than the one with the center crop. Manual annotation of the faces (that were not detected by Face++) resulted in further improvements. This feature augmentation scheme showed that having the face as foreground and the four other crops as background provides sucient representation to mimic the foreground and background contrast rule raised in [3]. Figure 15 shows some sample images that were correctly classified (true positive, true negative) and incorrectly classified (false positive, false negative).

Meanwhile, we observe that the GoogleNet model outperforms the Overfeat model for each feature augmentation approach (see Figure 16b). This is expected of a more contemporary DCNN architecture that is deeper and had previously yielded better results at the ILSVRC. Of course, this comes at the expense of heavier computational cost. Figure 8 shows the breakdown of classification performance into the the five parts as specified in [13]. There is no monopoly by any of the competing methods in the proposed scheme while the warping technique is noticeably poor.

Conventional single image warping produced quite poor results, much worse than Baseline B [13] which relies on an eclectic bunch of features that are both handcrafted and learned. It is important to point out that DCNN based methods do not required careful selection of features [3], or needing to eliminate features that are poorly correlated [13]. This itself gives an immense advantage towards methods that are able to learn good features on it own with minimal intervention.

5.2 Influence of small faces

We made an interesting observation that the AVA-Portraits dataset contains so-called "portrait" images with small faces inside (see Figure 19 for some samples). This may be caused by the fact that photo challenges in DPChallenge [14] are based on themes

that can be subjectively interpreted by participants (as they deemed fit), or that these photos could exemplify certain artistic qualities associated to portraiture. Hence, this is the reason why some photos are not captured closed-up, frontal, or from portrait-like angles.



Fig. 15. *Classification result of some images from the AVA-Portraits dataset using the GoogleNet-c5f-m method. The four quadrants show sample images based on their actual ground truth aesthetic class and predicted aesthetic class (positive side indicates high aesthetic quality, negative side indicates low aesthetic quality).*

We identified images with "small faces" by considering faces that have a bounding box area of smaller or equal to 49 pixels1. We investigated if small faces can possibly influence the overall aesthetics of these portrait images. The whole dataset, and likewise, each of the five parts, contain approximately 17% of small face portrait images (Figure 16). So that we can conclude that small face portrait images, distributed almost evenly across the five parts. A total of 1,388 small face portrait images were found while the remaining 6,445 images contain normal size faces (those not marked by Face++ were ignored). Two sets of experiments were conducted on both small and normal size faces using the GoogleNet model. The first experiment uses the standard single image warping ***'warp'****, while the second uses the* ***'c5f'*** *five crop with face scheme. We indicate the faces sizes with* ***'nf'*** *for normal size and* ***'sf'*** *for small size.*

Table 5 reports the accuracy of these evaluated methods for each of the five parts, and their respective mean accuracy and dispersion error. We observe that the photos with small faces performed slightly poorer than their normal face counterparts, with a larger spread of results. Considering that the proportion of small faces are evenly divided in all parts (Figure 16), the increase in dispersion error could mean that there are other non-face related aesthetic elements (such as image condition, background context,

artistic styles, etc.) found in these portrait images that could potentially influence their actual ratings. Small faces are also likely to not contribute any meaningful impression of beauty to the viewer, in contrast to the way how close-up portraits do.

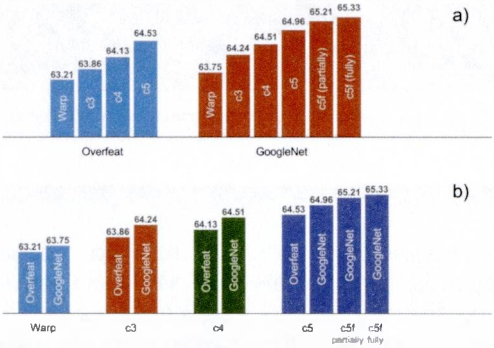


Fig. 16. Aesthetic classification accuracy grouped by (a) deep convolutional neural network models, and (b) type of feature augmentation scheme.

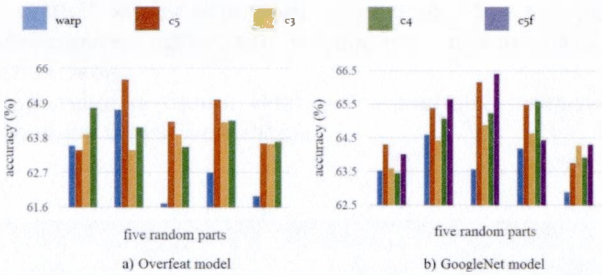


Fig. 17. Classification accuracy of the five random parts of the AVA-Portraits dataset using Overfeat and GoogleNet models.

Table 5. Comparison of classification performance between methods using small faces and normal faces on the GoogleNet-warp and GoogleNet-c5f methods.

Methods / Parts (%)	1st	2nd	3rd	4th	5th	Mean	Error
GoogleNet-warp-sf	65.10	62.96	63.12	62.16	59.91	62.50	2.59
GoogleNet-warp-nf	61.77	65.37	62.85	65.35	63.58	63.57	1.80
GoogleNet-c5f-sf	62.08	65.31	66.66	64.86	64.04	64.37	2.29
GoogleNet-c5f-nf	63.14	65.53	65.75	66.18	64.37	64.66	1.51

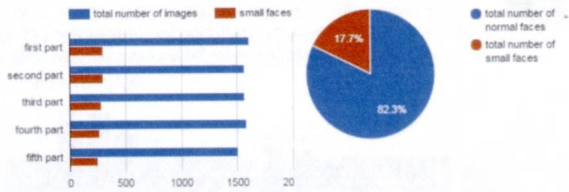


Fig. 18 *Proportion of small faces in AVA-Portraits dataset*

10. Қорытынды

- Қорытынды ретінде сіз не істегеніңізді және қандай нәтижелерге қол жеткізгеніңізді, сондай-ақ мәселені шешуде сіздің қрсқан үлесіңіз қандай болғанын қысқаша айтып өтуіңіз керек.
- Сондай-ақ, зерттеу процесінің ең қиын бөлігін (егер бар болса) сипаттап жазуыңызға болады.
- Зерттеу жұмысыңыздың соңында болашақта жетілдіруге арналған мәселелер тізімін беріп кеткен жөн.
- «Қорытынды» бөлімі «Нәтиже және оны талдау» бөлімінен бөлек жазылады.



Мысал:

Conclusion

In conclusion, we introduce a new compositional-based augmentation scheme for deep convolutional neural network (DCNN) models to evaluate the aesthetic quality of facial portraits. Photographic rules relating to composition principles are adapted for portraits with the purpose of enriching inputs to the network with more visually pleasing crops. This way, we are able to generate a deeplylearned augmented feature set that can improve classification. On the AVA-Portraits dataset (a subset of the AVA dataset first used in [13]) consisting of more than 10,000 portrait-like images, our approach demonstrated an improvement in



aesthetic classification over handcrafted baselines and standard DCNN models, achieving a top accuracy of 65.33%. In future, we intend to extend this idea by learning the compositional principles in facial portraits directly from the deep learning model itself. Also, we aim to explore other photo composition concepts besides the rule of thirds, such as visual balance and diagonal dominance.

11. Пайдаланылған әдебиеттер тізімі

Мақала соңында барлық пайдаланылған әдебиеттер тізімін міндетті түрде көрсету керек. Пайдаланылған әдебиеттерді қандай да бір тұрақты форматта жазған дұрыс. Осы кітаптың келесі бөлімдерінде LaTeX тілін және Google Scholar сервисін өту барысында пайдаланылған әдебиеттер тізімін қалай жазу керектігін көрсетеміз. Қазір біздің мысалға арналған пайдаланылған әдебиеттер тізімінің жазылу жолын ғана көрсетеміз.



Мысал:

References

1. Bakhshi, S., Shamma, D.A., Gilbert, E.: Faces engage us: Photos with faces attract more likes and comments on instagram. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems, ACM (2014) 965–974

2. Krages, B.: Photography: The art of composition. Skyhorse Publishing, Inc. (2012)

3. Khan, S.S., Vogel, D.: Evaluating visual aesthetics in photographic portraiture. In: Proceedings of the Eighth Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging, Eurographics Association (2012) 55–62

4. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. (2012) 1097–1105

5. Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229 (2013)
6. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1–9
7. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: European Conference on Computer Vision, Springer (2006) 288–301
8. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: 2011 International Conference on Computer Vision, IEEE (2011) 1784–1791
9. Murray, N., Marchesotti, L., Perronnin, F.: Ava: A large-scale database for aesthetic visual analysis. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE (2012) 2408–2415
10. Lu, X., Lin, Z., Jin, H., Yang, J., Wang, J.Z.: Rapid: rating pictorial aesthetics using deep learning. In: Proceedings of the 22nd ACM international conference on Multimedia, ACM (2014) 457–466
11. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. IEEE Trans. on Multimedia 15 (2013) 1930–1943
12. Tian, X., Dong, Z., Yang, K., Mei, T.: Query-dependent aesthetic model with deep learning for photo quality assessment. IEEE Transactions on Multimedia 17 (2015) 2035–2048
13. Redi, M., Rasiwasia, N., Aggarwal, G., Jaimes, A.: The beauty of capturing faces: Rating the quality of digital portraits. In: 11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG). Volume 1. (2015) 1–8
14. : Dpchallenge - a digital photography contest. (http://www.dpchallenge.com/) Retrieved: 26 August 2016.
15. Wang, Z., Dolcos, F., Beck, D., Chang, S., Huang, T.S.: Brain-inspired deep networks for image aesthetics assessment. arXiv preprint arXiv:1601.04155 (2016)
16. Lihua, G., Fudi, L.: Image aesthetic evaluation using paralleled deep convolution neural network. arXiv preprint arXiv:1505.05225 (2015)
17. Lienhard, A., Ladret, P., Caplier, A.: Low level features for quality assessment of facial images. In: 10th Int. Conf. on computer Vision Theory and Applications, VISAPP. (2015) 545–552
18. Kalayeh, M.M., Seifu, M., LaLanne, W., Shah, M.: How to take a good selfie? In: Proceedings of the 23rd ACM international conference on Multimedia (2015) 923–926
19. Obrador, P., Schmidt-Hackenberg, L., Oliver, N.: The role of image composition in image aesthetics. In: IEEE Int. Conf. on Image Processing. (2010) 3185–3188

20. : Face++: Leading Face Recognition on Cloud. (http://www.faceplusplus.com/) Retrieved: 26 August 2016.
21. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86 (1998) 2278–2324
22. : ImageNet Database and Challenge. (http://www.image-net.org/) Retrieved: 26 August 2016.
23. The sklearn-theano library. (https://sklearn-theano.github.io) Retrieved: 26 August 2016.
24. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531 (2014)
25. Child, J.: Studio photography: Essential skills. CRC Press (2013)

11. Қосымша

Егер сізде негізгі мақаланы жазу жұмысыңызға қосатын қосымшалар болса, онда одарды осы бөлімге енгізген дұрыс. Ол қосымшалар үлкен көлемдегі нәтижелер (үлкен кестелер немесе үлкен көлемдегі суреттер) немесе бөліскіңіз келетін кодтар болуы мүмкін. Бұлардан басқа қосқыңыз келген өзге де қосымшаларды осы бөлікке қоса аласыз.

Жоғарыда біз мысал ретінде көрсеткен мақаланың толық нұсқасы төменде берілді.

1

Aesthetic Evaluation of Facial Portraits Using Compositional Augmentation for Deep CNNs

2

Magzhan Kairanbay, John See, Lai-Kuan Wong
Center for Visual Computing, Faculty of Computing and Informatics
Multimedia University, 63100 Cyberjaya, Malaysia
magzhan.kairanbay@gmail.com,johnsee@mmu.edu.my,lkwong@mmu.edu.my

3

Abstract. Digital facial portrait photographs make up a massive portion of photos in the web. A number of methods for evaluating the aesthetics of photographs have been proposed recently. However, there have been a little work in the research community to address the aesthetics of targeted image domain, such as portraits. This paper introduces a new compositional-based augmentation scheme for aesthetic evaluation of portraits by well-known deep convolutional neural network (DCNN) models. We present a set of feature augmentation methods that take into account compositional photographic rules to ensure that the aesthetic in portraits are not hindered by standard transformations used for DCNN models. On a portrait subset of the large-scale AVA dataset, the proposed approach demonstrated a reasonable improvement in classification performance over the baseline and vanilla deep learning approaches.

4

Key words: Aesthetics, portraits, deep learning, Convolutional Neural Networks

5

1. Introduction

The ubiquity of smartphones and popularity of social media networks have contributed to an increasing number of photographs, particularly self-portraits or "selfies". A recent study by Bakhshi et al. [1] showed that photos with faces are more popular than other types of pictures. Nevertheless, casual photographers tend to produce many unprofessionally taken photos that are less aesthetically pleasing, and proceed to share them with other people in their circles. As such, an automatic assessment of the photo aesthetics would be useful to provide advice as to which of these photos should be retaken. This is also helpful for designers to select aesthetically pleasing photos for magazine covers, books, and news articles. However, different type of photos needs different type of photographic rules. Even, the most applicable rules, such as rule of thirds [2] could be applied differently for portraits and other type of photos [3].

Deep learning have been widely used in computer vision and natural language processing fields in recent years. In particular, deep convolutional neural network (DCNN) models have made revolutionary steps in object detection and recognition tasks [4–6]. The usage of graphic processing units (GPUs) have made it possible to handle large amounts of data, which directly provides better generalization and discrimination of models. Despite these technological advances and a growing research interest in computational aesthetics ever since the work of Datta et al. [7], very little work has been done in the specific domain of facial or portrait aesthetics, with a number of works treating the aesthetic problem generically [8–10] or categorically depending on its content [11] or query [12].

Motivated by these developments, we are inspired to improve existing deep learning techniques to better evaluate facial aesthetics. In this paper, we present a new compositional-based augmentation scheme for deep convolutional neural networks (DCNN) that can boost the aesthetic evaluation of facial portraits. Conventional sampling of sub-images (i.e. warping, center-crop, random crop) as receptive fields of a DCNN is rigid and can potentially distort the composition in portraits (Figure 1). As such, we exploit the rule of thirds composition rule to generate a feasible set of inputs to the network, which are in turn aggregated to form a deeply-learned augmented feature set. On a portrait-based dataset derived from the large-scale AVA dataset, our approach is able to surpass the performance of handcrafted and compositional-based features [3, 13] which have been a staple for defining beauty in portraits. The objective of this paper is to provide the methods of evaluating the portraits aesthetics using deep neural networks.

The following sections are organized as follows: Section 2



Fig. 1. Compositional-based cropping for DCNN inputs: (a) A portrait can be cropped aesthetically from power points defined by rule of thirds composition principle, (b) Full image warping (top) and center-cropping (bottom) can typically distort composition in portraits. (c) The four sub-images based on the power points.

reviews the current works in literature, Section 3 describes the dataset that was used, Section 4 elaborates on the proposed method in detail. Section 5 reports the experimental results with further discussion on them, while Section 6 concludes the paper.

2. Related work

The task of evaluating the aesthetic quality of photographs is becoming increasingly popular in recent years. As the notion of aesthetics is a subjective one, substantial human effort to rate photos is needed. Nonetheless, the creation of various datasets such as CUHKPQ [11] and AVA [9], from online photo communities such as Photo.net and DPChallenge [14], have made it possible to evaluate solutions with ground truth labels. Works from the early efforts of Datta et al. [7] to the recent use of popular descriptors [8] have concentrated on the extraction of hand-crafted visual features for aesthetic evaluation of photos. In [7], as many as 66 pre-determined features, depicting colourfulness, composition, textures and statistical properties were crafted while [8] went for generic representations such as SIFT and Color descriptors, supplemented by Bag-of-Words (BoW) and Fisher Vector (FV) encodings and spatial pyramids (SP).

Different forms of convolutional neural networks [15, 10, 16] have been applied to solve this task, with the intent of learning features in a more natural and unguided manner. We take particular interest in the work of Lu et al. [10], which generated two heterogeneous inputs to represent global and local cues respectively. Global views are represented by normalized inputs such as center-crop, warp and padding, while local views are represented by randomly-cropped inputs of smaller parts of the image. These views were used generically for any type of photo, and may not be suitable for photos with facial portraits. However, a majority of DCNN-based works showed scant attempts to consider compositional impact relating to photographic rules.

However, not many works focused on the aesthetics of photos from a specific domain, particularly in facial portraits. The works in [3, 13, 17] try to rate the aesthetic quality of face portraits and understand the visual features that define the beauty of digital portraits, while another recent work [18] attempted to answer the question of "how to take a good selfie?" from the perspective of image popularity.

Khan and Vogel [3] made a comparative study on the use of traditional features [7] versus their choice of features that are focused on spatial and color composition. Firstly, template based spatial composition features were constructed from three templates for power points, power lines and their combination which were modified from [19]. Secondly, highlight and shadow composition features were obtained by extracting the face illumination, background contrast, brightness and size of the face as features. Their results show that the proposed features are more informative, more concise and better equipped to rank portraiture composition. The proposed features are more accurate by 6% and smaller 9 times in feature length compared to traditional features. More importantly, the authors postulated that spatial composition is more important than highlight and shadow composition when identifying aesthetics.

Redi et al. [13] introduced their own features to describe the portrait aesthetics. These features have been derived from the following photographic dimensions: compositional rules (lighting, sharpness, traditional compositional features), semantics and scene content (object bank feature), basic quality metrics (noise, exposure quality, JPEG quality), portrait specific features (facial landmark sharpness, landmark statistics and face/background contrasts) and fuzzy properties (emotion, originality, memorability and uniqueness). Their feature analysis showed that race, gender and age of a person are uncorrelated with photographic beauty, while soft and hard image properties such as sharpness of facial landmarks, image contrast, exposure, uniqueness and originality are directly related to portrait aesthetics. The proposed method (after shrinking to a selected set of 12 positively-correlated features) outperformed the baseline method of [3] by 14.66% on the same data. However, using only face-related features without any selection, it performed better by only ~1%. Another work by Lienhart et al. [17] applied similar types of features on three small datasets, while also performing aesthetic score prediction by support vector regression. But their work does not have the similar rigor and extensiveness as

that of [13].

7

3. Dataset

The Aesthetic Visual Analysis (AVA) [9] dataset is the largest known dataset that is widely used to evaluate the aesthetic value of the images. The dataset contains more than 250,000 images from DPChallenge [14] with ground truth data such as human ratings ranging from 1 to 10 (with average 210 ratings per photo), with further meta-annotations from 66 textual semantics and 14 photographic styles from 72 different photo challenges. Following the work in [13], we select from the entire AVA dataset photos that are tagged with "Portraits" and also those whose challenge title contained the words Portrait, Portraiture or Portraits. Then, we filter further by detecting for photos that contained the presence of one or more faces with Face++ API [20]. A total of 10,141 images were obtained to form the portrait subset dataset. For better clarity, we refer to this subset as "AVA-Portraits" from this point onwards.

The same evaluation methodology as [13] was used. The dataset is randomly divided into five equal partitions. To ensure fair experimentation, we obtained the partition information from the authors. One partition will be used as test set and the rest as training set. Using 5-fold cross validation, the final accuracy is determined by averaging the minimum and maximum result of these five different cases. A portrait image is classified as aesthetically high if its aesthetic score is more than 5.5 (mean user score) [13]. Otherwise, the image is classified as aesthetically low.

8

4. Proposed approach

Convolutional neural networks (CNN) are biologically-inspired variants of the multilayer perceptron, dating back to the work of LeCun et al. [21]. However, it started to be increasingly popular only in 2012, after the work of Krizhevsky et al. [4] introduced an eight-layer deep CNN network, winning the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. From then on, convolutional neural network models have shown unrivaled performance in ILSVRC each year. In ILSVRC, models are trained and tested on the ImageNet dataset [22], which consists of more than 14 million of images with 1,000 classes. The classes comprise of a large variety of objects such as car, motocycle, person, cats, etc. Realistically, an immense amount of computational power is necessary to be able to train weights in these deep networks at a reasonable amount of time. Thus, pre-trained models are commonly available for application to different related classification problems. However, in view of the different number of classes in other problems, the number of nodes in the last softmax layer should be replaced with that equivalent to the number of classes. In our case, the softmax layers are 2x1 as it is a binary classification problem (good aesthetics, bad aesthetics) (See Figure 2).

There are a number of popular pre-trained models available for public use. In this work, we use the Overfeat [5] (winner of the localization task in ILSVRC 2013) and GoogleNet [6] (winner of the classification task in ILSVRC 2014). For each of these models, the number of nodes in last (softmax) layer is reduced from 1000 to 2, and the weights are fine-tuned with the new training data. The second last layer of the network is then taken as the feature vector for classification by a softmax classifier (or multi-class binary logistic regression). We use the Sklearntheano [23] library, an open source library on the Theano framework for the implementation of this research.

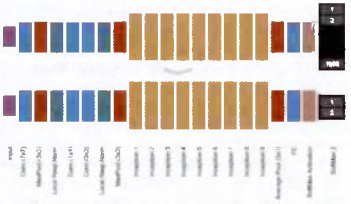


Fig. 2. Changing the number of nodes in the last layer of GoogleNet.

4.1 Feature augmentation using photographic rules

Feature augmentation is a process intended to increase the feature vector length at a certain layer by aggregating features obtained from various transformed versions of the input image. We distinguished this from *data augmentation*, which is a well-known technique [24, 4] used especially for learning deep representations by generating additional number of training samples. By augmentation we mean perturbing an image by transformations (e.g. cropping, flipping, or a combination), leaving the underlying class unchanged.

In our work, we propose a feature augmentation scheme based on compositional photographic rules. This is motivated by the fact that typical image transformations such as cropping and flipping do not make aesthetic

sense when it comes to face portraits. They may aggravate the actual composition of the portrait since a majority of cropping are done randomly or without applying compositional principles, while warping can potentially distort the face aspect ratio. The rule of thirds [19, 2, 25] is one of the popular photographic rules which divides the image to nine parts by 3 x 3 (see Figure 3). The four intersections of the rule of thirds lines are called power points. According to this popular rule, a photo is considered aesthetically appealing if the object of interest is located near one of these four power points.

We design a number of compositional cropping arrangements that are motivated by photographic rules associated to rule of thirds. Firstly, the baseline transformation involves warping the image to a constant width and height (corresponding to the model) and passing it as a single input to the DCNN model. We denote this in our experimental reporting as **'warp'**.

According to [3], the face, which is the main subject of interest in the portrait, should be located at the top power line by keeping the eyes in the upper third of the image area. Meanwhile, the symmetrical composition rule states that it is aesthetically pleasing to center the face in the image frame. Inspired by these photographic composition rules, we modify the rule of thirds for portraits by taking two crops centered on the top two power points (left and right) and a center crop between these two crops (see Figure 3a). All three crops are passed to the DCNN model as individual inputs to generate feature vectors that are then concatenated. We denote this technique as **'c3'**. The third technique (denoted as **'c4'**) uses four different sub-images of constant height and width, center cropped from four power points based on the rule of thirds (see Figure 3b). The augmented feature vector is formed by concatenating the four feature vectors generated from each crop.



Fig. 3. a) Three-crop by rule of thirds, b) Four-crop by rule of thirds.

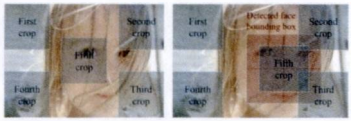


Fig. 4. a) Five-crop: four crops from four different corners + center crop) b) Five-crop: four crops from four corners + face crop.

Extending on to five-crop method, we propose two different variants. The first (denoted as **'c5'**) takes four sub-images from the four corners of the image, and the center crop as the fifth. Since the fifth crop provides information from the middle of the photo, the use of four crops from the corners yields better coverage than using crops centered at all four power points. The second variant (denoted as **'c5f'**) crops the face region as the fifth crop. Face++ [20] API was used to detect the location of the center coordinate, and a bounding box that defines the sub-image is cropped and squashed to a constant width and height. Figure 5 shows some sample face regions detected by Face++ API and used as the fifth crop for method **'c5f'**. Similar to previous schemes, DCNN features resulting from all five crops are concatenated. Due to slight unreliability of the Face++ API, only about 77.24% of all faces from the AVA-Portraits dataset can be detected. As such, the affected images fall back on the center crop as the fifth crop. Alternatively, we also manually annotated the faces in the remaining affected images. We denote this as **'c5f-m'**.



For all compositional cropping methods, we crop the sub-images based on a fixed width and height of 256 pixels, which are subsequently resized to receptive field sizes of their respective networks i.e. 234 for Overfeat, 224 for GoogleNet, both in RGB color space with zero mean.

Table 1. Aesthetic classification accuracy and dispersion error of the proposed methods compared against baseline methods.

Methods	Accuracy (%)	Error ± (%)
Baseline A [3]	54.585	1.25
Overfeat-warp	63.211	1.48
GoogleNet-warp	63.751	0.85
Overfeat-c3	63.861	0.44
Overfeat - c4	64.133	0.62
Baseline B [13]	64.240	1.76
GoogleNet-c3	64.244	0.64
GoogleNet-c4	64.519	1.06
Overfeat-c5	64.536	1.11
GoogleNet-c5	64.962	1.20
GoogleNet-c5f	**65.216**	1.20
GoogleNet-c5f-m	**65.330**	1.01

5 Results and Discussions

5.1 Overall results

We define two baselines for the purpose of comparisons– one by [3] which relies on traditional spatial and color composition features (**'Baseline A'**), and another [13] which uses a discriminative set of visual features (**'Baseline B'**). Two DCNN models were utilized in our experiments: the 8-layer **Overfeat** model and the 22-layer very deep **GoogleNet** model.

Table 1 shows the proposed and baseline approaches on the AVA-Portraits dataset. We achieve the best result reported thus far for this dataset, of 65.33%. From the feature augmentation results, we can see that the aesthetic classification accuracy generally increases as more crops were added to augment the DCNN-learned features (see Figure 7a). The best performing scheme uses five crops, with the variant that takes a face as fifth crop marginally higher than the one with the center crop. Manual annotation of the faces (that were not detected by Face++) resulted in further improvements. This feature augmentation scheme showed that having the face as foreground and the four other crops as background provides sucient representation to mimic the foreground and background contrast rule raised in [3]. Figure 6 shows some sample images that were correctly classified (true positive, true negative) and incorrectly classified (false positive, false negative).

Meanwhile, we observe that the GoogleNet model outperforms the Overfeat model for each feature augmentation approach (see Figure 7b). This is expected of a more contemporary DCNN architecture that is deeper and had previously yielded better results at the ILSVRC. Of course, this comes at the expense of heavier computational cost. Figure 8 shows the breakdown of classification performance into the the five parts as specified in [13]. There is no monopoly by any of the competing methods in the proposed scheme while the warping technique is noticeably poor.



Fig. 6. Classification result of some images from the AVA-Portraits dataset using the GoogleNet-c5f-m method. The four quadrants show sample images based on their actual ground truth aesthetic class and predicted aesthetic class (positive side indicates high aesthetic quality, negative side indicates low aesthetic quality).

Conventional single image warping produced quite poor results, much worser than Baseline B [13] which relies on an eclectic bunch of features that are both handcrafted and learned. It is important to point out that DCNN based methods do not required careful selection of features [3], or needing to eliminate features that are poorly correlated [13]. This itself gives an immense advantage towards methods that are able to learn good features on it own with minimal intervention.

5.2 Influence of small faces

We made an interesting observation that the AVA-Portraits dataset contains so-called "portrait" images with small faces inside (see Figure 10 for some samples). This may be caused by the fact that photo challenges in DPChallenge [14] are based on themes that can be subjectively interpreted by participants (as they deemed fit), or that these photos could exemplify certain artistic qualities associated to portraiture. Hence, this is the reason why some photos are not captured closed-up, frontal, or from portrait-like angles.

We identified images with "small faces" by considering faces that have a bounding box area of smaller or equal to 49 pixels1. We investigated if small faces can possibly influence the overall aesthetics of these portrait images. The whole dataset, and likewise, each of the five parts, contain approximately 17% of small face portrait images (Figure 9). So that we can conclude that small face portrait images, distributed almost evenly across the five parts. A total of 1,388 small face portrait images were found while the remaining 6,445 images contain normal size faces (those not marked by Face++ were ignored). Two sets of experiments were conducted on both small and normal size faces using the GoogleNet model. The first experiment uses the standard single image warping '**warp**', while the second uses the **'c5f'** five crop with face scheme. We indicate the faces sizes with **'nf'** for normal size and **'sf'** for small size.

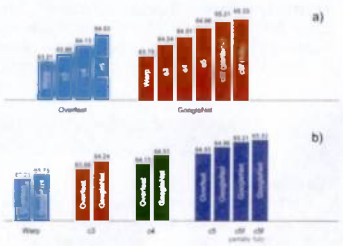


Fig. 7. Aesthetic classification accuracy grouped by (a) deep convolutional neural network models, and (b) type of feature augmentation scheme.

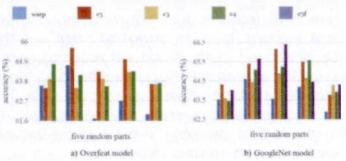


Fig. 8. Classification accuracy of the five random parts of the AVA-Portraits dataset using Overfeat and GoogleNet models.

Table 2. Comparison of classification performance between methods using small faces and normal faces on the GoogleNet-warp and GoogleNet-c5f methods.

Methods / Parts (%)	1st	2nd	3rd	4th	5th	Mean	Error
GoogleNet-warp-sf	65.10	62.96	63.12	62.16	59.91	62.50	2.59
GoogleNet-warp-nf	61.77	65.37	62.85	65.35	63.58	63.57	1.80
GoogleNet-c5f-sf	62.08	65.31	66.66	64.86	64.04	64.37	2.29
GoogleNet-c5f-nf	63.14	65.53	65.75	66.18	64.37	64.66	1.51

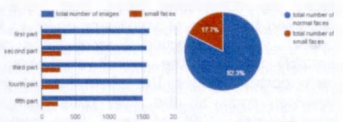


Fig. 9. Proportion of small faces in AVA-Portraits dataset

Table 2 reports the accuracy of these evaluated methods for each of the five parts, and their respective mean accuracy and dispersion error. We observe that the photos with small faces performed slightly poorer than their normal face counterparts, with a larger spread of results. Considering that the proportion of small faces are evenly divided in all parts (Figure 9), the increase in dispersion error could mean that there are other non-face related aesthetic elements (such as image condition, background context, artistic styles, etc.) found in these portrait images that could potentially influence their actual ratings. Small faces are also likely to not contribute any meaningful impression of beauty to the viewer, in contrast to the way how close-up portraits do.

6 Conclusion

In conclusion, we introduce a new compositional-based augmentation scheme for deep convolutional neural network (DCNN) models to evaluate the aesthetic quality of facial portraits. Photographic rules relating to composition principles are adapted for portraits with the purpose of enriching inputs to the network with more visually pleasing crops. This way, we are able to generate a deeplylearned augmented feature set that can improve classification. On the AVA-Portraits dataset (a subset of the AVA dataset first used in [13]) consisting of more than 10,000 portrait-like images, our approach demonstrated an improvement in aesthetic classification over handcrafted baselines and standard DCNN models, achieving a top accuracy of 65.33%. In future, we intend to extend this idea by learning the compositional principles in facial portraits directly from the deep learning model itself. Also, we aim to explore other photo composition concepts besides the rule of thirds, such as visual balance and diagonal dominance.

References

1. Bakhshi, S., Shamma, D.A., Gilbert, E.: Faces engage us: Photos with faces attract more likes and comments on instagram. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems. ACM (2014) 965–974
2. Krages, B.: Photography: The art of composition. Skyhorse Publishing, Inc. (2012)
3. Khan, S.S., Vogel, D.: Evaluating visual aesthetics in photographic portraiture. In: Proceedings of the Eighth Annual Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging. Eurographics Association (2012) 55–62
4. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. (2012) 1097–1105
5. Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229 (2013)
6. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1–9
7. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: European Conference on Computer Vision, Springer (2006) 288–301
8. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: 2011 International Conference on Computer Vision, IEEE (2011) 1784–1791
9. Murray, N., Marchesotti, L., Perronnin, F.: Ava: A large-scale database for aesthetic visual analysis. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE (2012) 2408–2415
10. Lu, X., Lin, Z., Jin, H., Yang, J., Wang, J.Z.: Rapid: rating pictorial aesthetics using deep learning. In: Proceedings of the 22nd ACM international conference on Multimedia, ACM (2014) 457–466
11. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. IEEE Trans. on Multimedia 15 (2013) 1930–1943
12. Tian, X., Dong, Z., Yang, K., Mei, T.: Query-dependent aesthetic model with deep learning for photo quality assessment. IEEE

Transactions on Multimedia 17 (2015) 2035–2048
13. Redi, M., Rasiwasia, N., Aggarwal, G., Jaimes, A.: The beauty of capturing faces: Rating the quality of digital portraits. In: 11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG). Volume 1. (2015) 1–8
14. Dpchallenge - a digital photography contest. (http://www.dpchallenge.com/) Retrieved: 26 August 2016.
15. Wang, Z., Dolcos, F., Beck, D., Chang, S., Huang, T.S.: Brain-inspired deep networks for image aesthetics assessment. arXiv preprint arXiv:1601.04155 (2016)
16. Lihua, G., Fudi, L.: Image aesthetic evaluation using paralleled deep convolution neural network. arXiv preprint arXiv:1505.05225 (2015)
17. Lienhard, A., Ladret, P., Caplier, A.: Low level features for quality assessment of facial images. In: 10th Int. Conf. on computer Vision Theory and Applications. VISAPP. (2015) 545–552
18. Kalayeh, M.M., Seifu, M., LaLanne, W., Shah, M.: How to take a good selfie? In: Proceedings of the 23rd ACM international conference on Multimedia (2015) 923–926
19. Obrador, P., Schmidt-Hackenberg, L., Oliver, N.: The role of image composition in image aesthetics. In: IEEE Int. Conf. on Image Processing. (2010) 3185–3188
20. Face++ Leading Face Recognition on Cloud. (http://www.faceplusplus.com/) Retrieved: 26 August 2016.
21. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86 (1998) 2278–2324
22. ImageNet Database and Challenge. (http://www.image-net.org/) Retrieved: 26 August 2016.
23. The sklearn-theano library. (https://sklearn-theano.github.io) Retrieved: 26 August 2016.
24. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531 (2014)
25. Child, J.: Studio photography: Essential skills. CRC Press (2013)

13. LaTeX-ті қолдану әдістемесі

Мақала жазу барысында түрлі жазу редакторын қолдана аласыз. Ол жазу редакторларының ішіндегі ең танымалдары Microsoft Word және LaTeX. Мақалаларды Microsoft Word-қа қарағанда LaTeX-пен жазу тиімдірек. Себебі LaTeX арқылы жазылған мақаланың құрылымын бір рет дайындап алып, кейін соған өзіңіздің мақала мәтінін енгізу арқылы форматтан ауытқымайтыныңызға сенімді бола аласыз. Сонымен бірге LaTeX арқылы жазылған мақала сәндірек әрі ұқыптырақ болып келеді. LaTeX-пен жазу барысында мақала құрылымын LaTeX коды арқылы дайындап алғаныңыз керек. Жалпы, LaTeX-пен жазу барысында мақаланы LaTeX командаларын қолдану арқылы жазасыз. LaTeX кодын жазғаннан кейін ол кодты компиляциядан өткізіп, керекті мақаланы шығарып аламыз. LaTeX коды мен мақала төмендегі суретте көрсетілген (19-сурет). Суретте көрсетілген LaTeX кодын жазу редакторы Overleaf деп аталады. Көріп отырғаныңыздай, оң жақта LaTeX коды, ал сол жақта мақала көрсетілген. Overleaf арқылы мақала жазудың бір артықшылығы — бірнеше адамның бір мақаланы өзгерту арқылы жұмыс істей алатында. Яғни бірнеше адам мақала жазу барысында олардың әр-қайсысы өз бөлігін жазып, бір мақаланы осылайша толықтыра алады. Жиі кездесетін LaTeX командаларын төмендегі документ арқылы көре аласыз:
https://wch.github.io/latexsheet/latexsheet.pdf

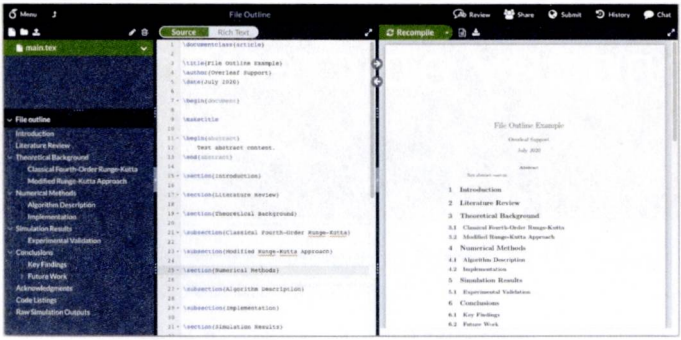


Fig. 19. *Sample Latex code with Overleaf service*

14. Google Scholar-ді қолдану әдістемесі

Ғылыми-зерттеудегі әдеби шолу жасаудың барысында басқа да ғылыми-зерттеулерге талдау жасаған керек болады. Ол ғылыми жұмыстарға керекті материалдарды іздестіру барысында Google Scholar сайтын қолданған дұрыс. Google Scholar-ға https://scholar.google.com сайты арқылы кіруге болады. Google Scholar-дың басты беті мынадай болып келеді.

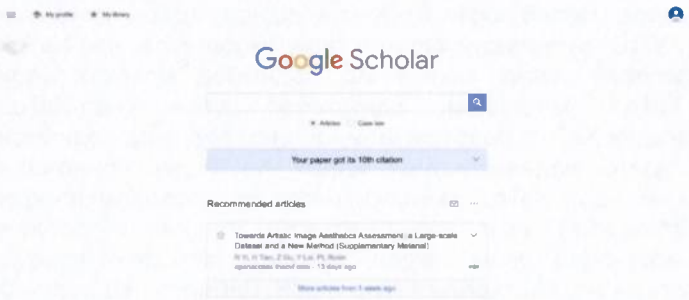


Fig. 20 *Main page of Google Scholar*

Яғни осы беттегі іздестірілетін жолға өзіңізге керекті кілтті сөздерді жазып, басқа да ғылыми жұмыстарды іздей аласыз. Мысалы, суреттің әдемілігіне байланысты ғылыми-зерттеулерді іздестірген кезде, сол жерге "photo aesthetics" деп жазуымызға болады. Осылай жазғаннан кейін мынадай мақалалар пайда болады (21-сурет).

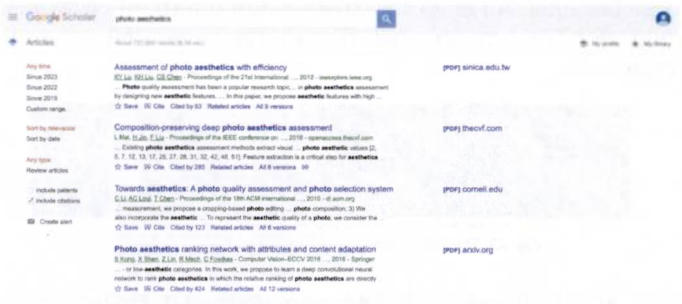


***Fig. 21**. The result of the request from Google Scholar*

Осы мақалалар тізіміндегі әр мақалаға кіріп, оны оқи аламыз және осы парақшаның сол жағындағы мақалаларды фильтрлей аламыз. Ол фильтрлер көбінше уақыт бойынша болады. Ал бірінші мақаланы оқып, кейін сол мақаланы өз ғылыми жұмысымызға қосқымыз келсе, онда оған сілтеме жасау керек. Мысалы, бірінші мақалаға сілтеме жасау үшін **"cite"** батырмасын басу керек. Бұл батырманы басқаннан кейін төмендегідей сілтеме жасау жолдары көрсетіледі (22-сурет).

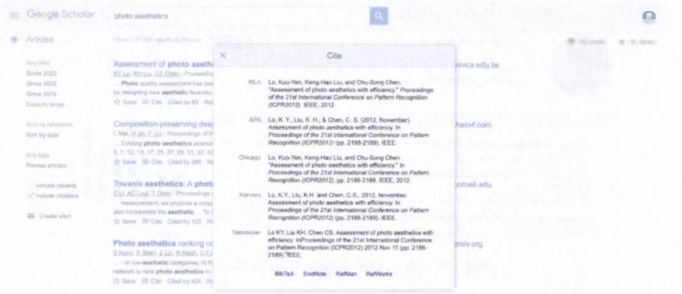


***Fig. 22**. The references from Google Scholar*

Бұл сілтемені көрсету жолдарында неше түрлі форматтар келтірілген. Өзіңізге керекті формат бойынша сілтемені көшіріп алып, мақалаңызға қосса-

ңыз болады. Ал егер сіз LaTeX-ті қолданатын болсаңыз, онда BibTeX батырмасын басып, төмендегідей кодқа ие боласыз.

```
@inproceedings{lo2012assessment,
  title={Assessment of photo aesthetics with efficiency},
  author={Lo, Kuo-Yen and Liu, Kang-Hao and Chen, Chu-Song},
  booktitle={Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)},
  pages={2186--2189},
  year={2012},
  organization={IEEE}
}
```

***Fig. 23**. The bibTex string generated from Google Scholar*

Бұл кодты өзіңіздің LaTeX құжатыңызға енгізіп, кейін осы мақалаға сілтеме көрсету үшін \cite{lo2012assessment} деп жазып қою қажет, мұндағы lo2012assessment — мақаланың аты. Осылайша, LaTeX-те басқа мақалаларға сілтеме жасай аласыз.

Google Scholar-дың негізгі беттіндегі "My Profile" батырмасын бассаңыз, өзіңіздің мақалаларыңызды және сіздің мақалаларыңызға неше рет сілтеме жасағанын көре аласыз. Бұл бет мына төмендегідей болады.

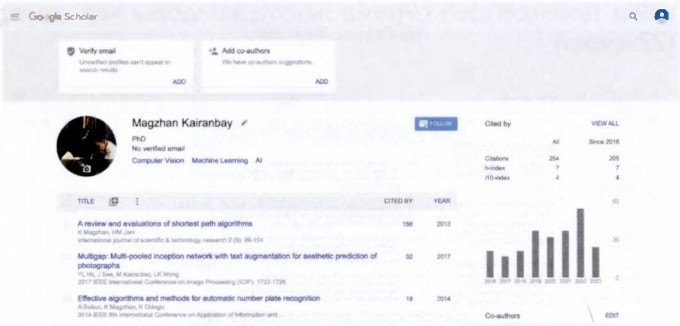


***Fig. 24**. "My profile" page of Google Scholar*

Осы бетте әр мақалаңызға жасалған сілтеме саны және әр жылдағы барлық сілтемелер санын көре аласыз. Сілтеме санын баса отырып, сіздің мақалаңызды пайдаланған басқа мақалаларды көре аласыз. Бірінші мақалаға арналған сілтеме жасалған ғылыми жұмыстар түрі төмендегідей.

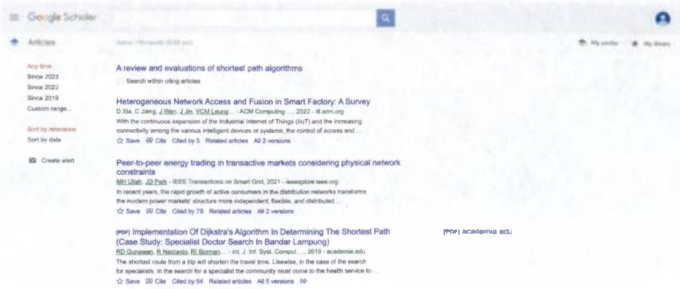


***Fig. 25**. Papers that have cited your paper*

Google Scholar-ді қолдану барысында әр автордың мақалалар тізімін көре аласыз. Болашақта сол авторлардың ғылыми-зерттеу жұмыстарымен танысу үшін, сол авторға "жазылсаңыз" да болады. Жазылу арқылы сол автор жазған жаңа мақалалар сіздің поштаңыз арқылы өзіңізді хабардар етіп отырады. Осылайша, сол автордың жаңа мақалаларын дер кезінде оқып тұра аласыз.

15. Презентация мен постер жасау жолдары

Конференцияларға қатысқанда ғылыми жұмысыңызды басқаларға таныстыру үшін презентация немесе постер жасайсыз. Постер дегеніміз — сіздің ғылыми жұмысыңызды таныстыратын бір беттен тұратын инфографика.

Презентация мен постер жасау барысында графикалық элементтерді дұрыс қолдану керек. Ол графикалық элементтер төмендегідей болуы мүмкін:

- Суреттердің қолданылуы.
- Мәтіннің қолданылуы.
- Инфографикалар.
- Түстер.
- Туралау.

Осы кітапта бұл элементтердің әрқайсысына бір-бірден мысал келтіре отырып түсіндіретін боламыз. Жалпы, презентация слайдтары мен постерді дайындау барысында сөзді аз, визуалды элементтерді көбірек қолдануға тырысыңыз. Слайд немесе постер дизайнына қалай болса солай қарамай, оған барынша дайындықпен келуге тырысқан жөн. Қарапайым дизайндар құрмай, оған шығармашылық ойларыңызды барынша көбірек қосқаныңыз дұрыс.

Слайдтарыңызда көлемі ұзақ мәтінді қолданудан аулақ болыңыз. Көлемі ұзақ мәтінмен сипаттама берудің орнына инфографиканы қолданып, презентацияны сол инфографиканы сөзбен түсіндіре отырып жасай аласыз. Немесе сөз болып отырған оқиғаға байланысты бір сурет қойып,

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

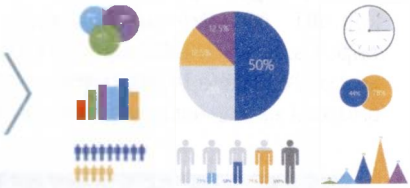


Fig. 26. Better to use less words and more visual infographics

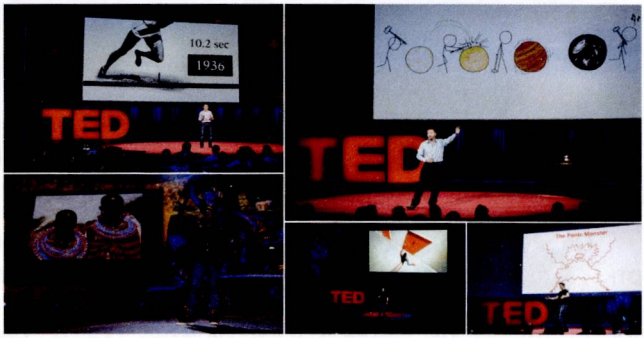


Fig. 27. Speakers of TED Talks practice having one visual and explain them during the talk

презентацияны сөзбен сипаттай отырып жасаған дұрыс. Осындай презентация жасау жолын әйгілі TED Talks спикерінен жиі көреміз.

Слайдтарды немесе постерлерді дайындау барысында әртүрлі бағдарламалар қолданылады. Олардың жиі кездесетіні — Power Point және Google Slides. Google Slides пен Power Point — бір-біріне өте ұқсас бағдарламалар. Оларда кәсіби презентация мен постер дайындауға арналған барлық функциялар бар. Презентацияның немесе постердің дизайнын жасағанда барынша уақыт пен күш-жігерді барынша

көп жұмсап, барынша шығармашылыққа баса көңіл бөлу керек. Шығармашылықты арттыру үшін түрлі дизайндарға шолу жасап, олардан өзіңізге ұнаған керекті дизайн элементтерін тауып, жаттап алуға тырысыңыз. Мысалы, төмендегі дизайндарға шолу жасап, өзіңізге ұнаған элементтерді жаттап, болашақта оларды өз жұмысыңызға қосып алуға тырысыңыз.

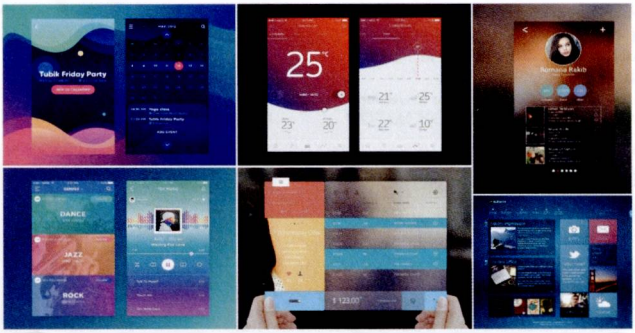


Fig. 28. Review to different UI / UX designs

Өз жұмысыңызда суреттерді қою кезінде олардың артқы фоны түссіз болуы керек. Егер сіздің слайдыңыздың фоны бір түсті болып, ал қойылған суреттің фоны ақ немесе өзге бір түс болса, онда мұндай визуалды көрініс ұсқынсыз болады. Сол себепті бұл олқылықты түзеу үшін, қолданылып жатқан суреттің артқы фоны түссіз болғаны керек. Мысалы, төмендегі суретке жүгінсеңіз болады.

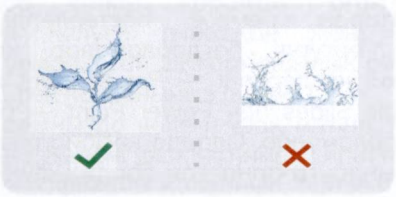


Fig. 29. Better to have transparent background rather than solid color

Сонымен бірге дизайнды бастамас бұрын интернеттен дизайн саласындағы трендтерге, стильдерге шолу жасап, сол тренд және стильдерді қолдануға тырысыңыз. Мысалы, төмендегі суретте дизайн стильдері көрсетіліп, олардың қайсысын қолдану немесе қолданбау туралы нұсқаулықтар берілген.



***Fig. 30**. Review to latest trends in UI / UX*

Суреттерді Google Slides-қа қою барысында түрлі жолдарды қолдануға болады. Олардың жиі қолданылатындары сілтеме және оларды жүктеу арқылы болады.

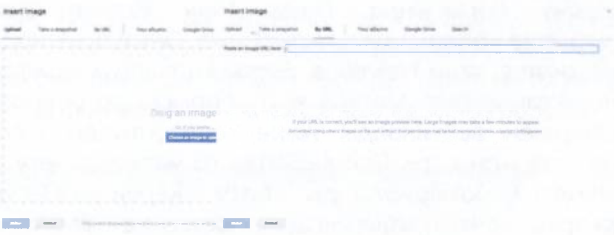


***Fig. 31**. Inserting images in Google Slides*

Инфографиканы әртүрлі формаларды қолдана отырып құрастыруға болады. Ол формаларға әрқандай түс беріп, олардың қырын және қырының қалыңдығын өзгерте аласыз. Бұл өзгерістерді формаға енгізу арқылы өз формаларыңызды әп-әдемі кейіпке келтіре аласыз. Google Slides-та қарапайым шеңбер құрып, сол шеңбердің түсін ауыстыру жолы мына төмендегідей болады.

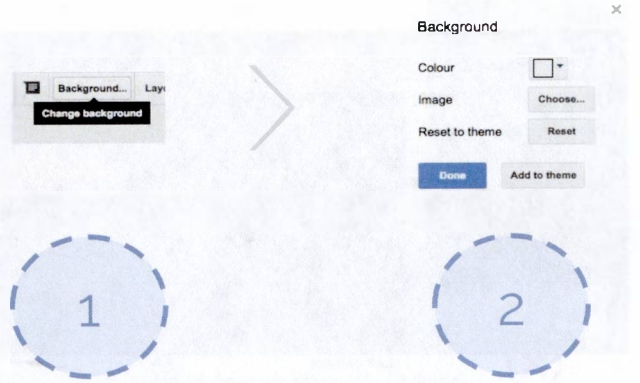


Fig. 32. Changing the shapes border and its color

Сонымен қатар суреттерді де түрліше формада кесуге болады. Мысалы, профиль суретін әдемілеу үшін суретті кесіп, қырын өзгертсек, сурет қарапайым төртбұрыш формасынан басқа бір әдемі формаға өзгереді. Google Slides-та бұл процессті іске асыру жолы төмендегідей (33-сурет).

Постер немесе слайдтарыңызда мәтіндерді қолдану барысында Times New Roman сияқты шрифтерді қолданбауға тырысыңыз. Оның орнына Sans Serif типтес, яғни Helvetica, Segoe UI сияқты шрифтерді қолданған дұрыс. Мәтінді жазу барысында оны парақ шеттерінен вертикалды және горизонталды жолмен түзеу өте маңызды. Google Slides-та мәтінді түзету үшін төмендегі батырмаларды басу керек (33-сурет). Жоғарыда айтып өткеніміздей, постер мен слайдтарда мәтінді аз қолдануға тырысыңыз. Сонымен бірге мәтін

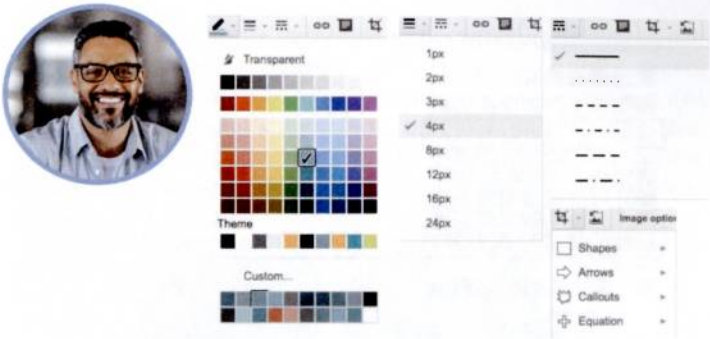


Fig. 33. Making the profile photo using Google Slides

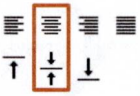


Fig. 34. Aligning the text borders vertically and horizontally

өлшемдерін үлкен қылуға тырысыңыз. Мәтіндерді Google Slides арқылы енгізіп, әрқандай түрлендірулер жасау үшін төмендегі батырмаларды бассуыңызға болады.

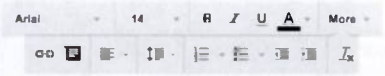


Fig. 35. Text tools in Google Slides

Мәтін көзге айқын көрінуі үшін фонның түсін жарық түстерге, ал мәтіннің түсін қараңғы түстерге бояңыз. Немесе фонды қараңғы түске, ал мәтінді жарық түске боясаңыз да болады. Google Slides-те төмендегі түстерді фон мен мәтінге қолданып көрсеңіз болады.

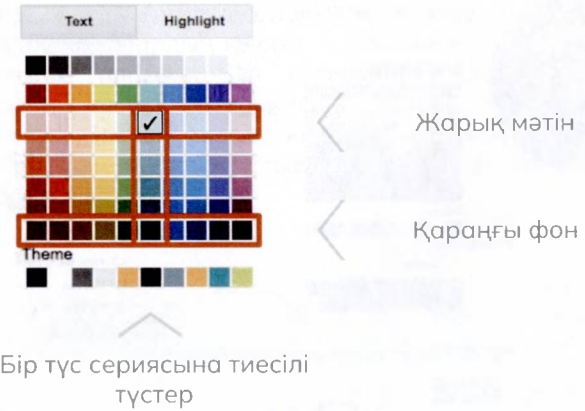


Fig. 36. Usage of different colors for foreground and background in Google Slides

Жоғардағы нұсқаулықтарды қолдана отырып жасалған бірнеше постердің мысалдарын көрсетейік.

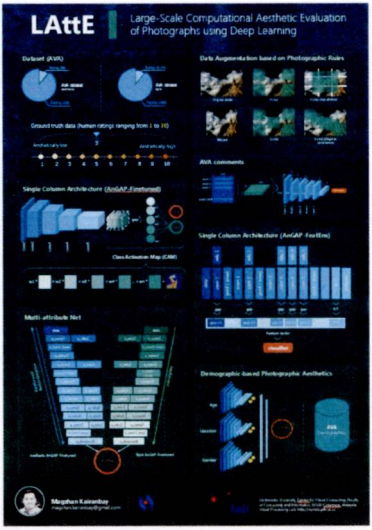


Fig. 37. Example of poster using Google Slides

Инфографикаларды құру өте қиын әрі күрделі процесс болып көрінеді. Бірақ бұл мүлдем олай емес. Инфографикаларды қарапайым геометрикалық фигуралар немесе формалар арқылы құруға болады. Алайда, бұл процессте шыдамдылық пен қажырлы еңбек керек. Қарапайым фигуралар мынадай болуы мүмкін.

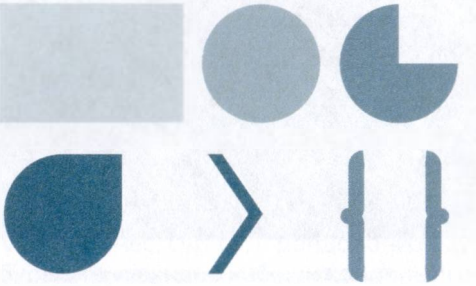


***Fig. 38**. Shapes in Google Slides*

Бұл фигураларды қолдану барысында бір түс сериясына тиесілі түстерді пайдалануға болады. Google Slides-та мына төмендегі түстерді қолдана аласыз.

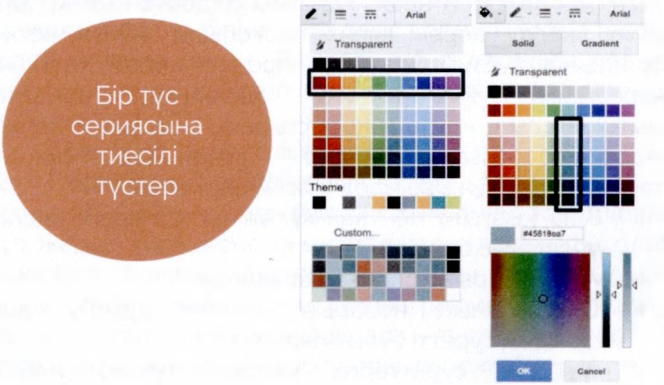


***Fig. 39**. Usage of colors in Google Slides*

Осы нұсқаулықтарға негізделе отырып жасалған инфографика мысалдары мынадай.

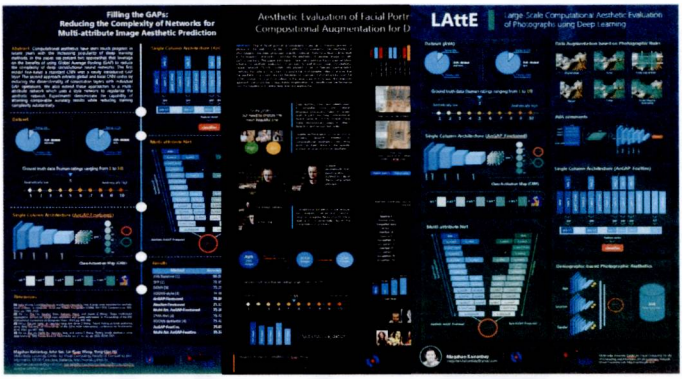


***Fig. 40**. Example of posters using Google Slides*

Кей жағдайда бір түспен анықталатын элементтерді көлеңке арқылы бөліп таста болады. Элементтерді осылайша бөлу әдісі жоғарыдағы мысалдарда қолданылды. Әдемі инфографикаларды салып шығу үшін, интернеттегі неше түрлі мысалдарға көз жүгіртіп, шолу жасау керек екенін жоғарыда айтып кеткен болатынбыз. Енді бұл шолу процесі қалай жүретінін көрсетейік. Жоғарыдағы дизайнға арналған мысалдарды қайта қарастырып, оларға талдау жасайық. Бізге берілген дизайндар арқылы төмендегідей тұжырымдар жасай аламыз:

- Қою қызғылт пен көк күлгін түстер өзара жақсы үйлесімде болады.
- Ақ түс барынша аз қолданылған.
- Оқырманның назарын мәтінге аудару үшін, фондық суретті бұлыңғыр ету.
- Фондық суреттерге мөлдір (тұнық) қабаты пайдаланып, оларды солғын көрсететіндей ету.
- Пішіндердің бұрышын қисықтау ету.
- Қисық пішіндес фигураларды қолдану.

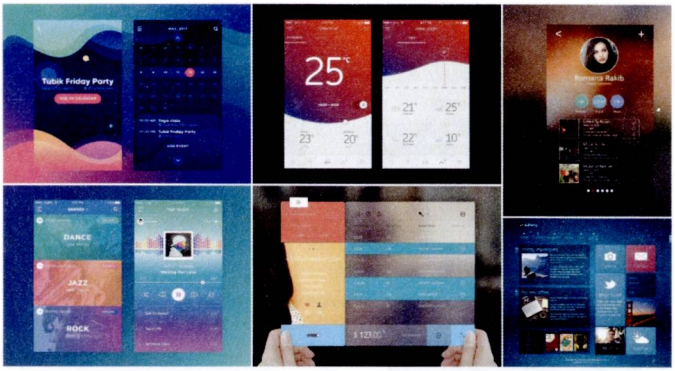


Fig. 37. Example of posters using Google Slides

- Мәтін қаріпі ақ болғаны дұрыс.
- Фон үшін ақ емес, сұр түсті пайдаланған дұрыс.
- Төртбұрышты пішіндер көбірек қолданылған
- Негізгі мәтіннің барлығы бас әріппен жазылған, қосымша мәтін кішірек және түсі ашық емес.

Инфографикаларды құру барысында түстердің дұрыс қолданылуы өте маңызды. Кей жағдайда бір түс сериясына тиесілі түстерді қолдану керек болса, басқа бір жағдайда бір-бірінен өте алыс қарама-қарсы түстер қолданылғаны абзал. Түстерді және олардың кодын ұсынатын түрлі сайттар бар. Солардың кейбіреулері мынадай:

1. **Colorhunt**, https://colorhunt.co.
2. **Flatuicolors**, https://flatuicolors.com.

Google Slides-та түстер былайша жолмен таңдалады. Бұл жерде түсті және оның мөлдірлік дәрежесін атап көрсетуге болады. Түстердің орынды қолданылуын дұрыс түсіну үшін басқа инфографикаларға шолу жасап, сол инфографикада түстердің қалай қолданылуына талдау жасау арқылы оны біліп, жаттап алуға тырысу керек. Мысалы, төмендегі инфографикада қолданылатын түстер және олардың алдыңғы және артқы фондағы қолданыстары көрсетілген.

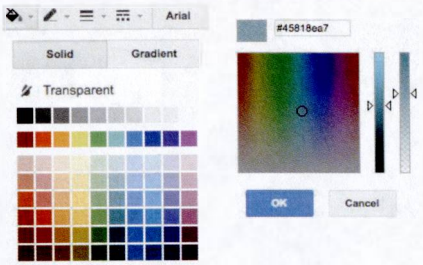


Fig. 38. Usage of colors in Google Slides

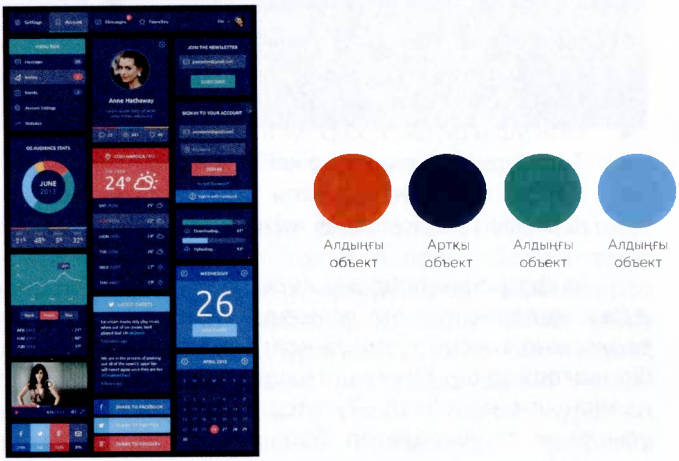


Fig. 39. Usage of colors in background and foreground

Дизайн жасау барысында элементтерді туралау өте маңызды. Пифагордың айтуы бойынша «үйлесімділік» және «симметрия» — сұлулық теориясының негізі. Сол себепті әдемі слайд немесе постер жасау үшін барлығы (суреттер, мәтіндер, фигуралар парақ бетімен салыстырғандағы орналасуы) симметриялы болу керек. Яғни қандай да бір элементті парақ бетіне қойғанда, сол элемент парақтың оң жақ, сол жақ және жоғарғы, төменгі деңгейлерінен бірдей қашықтықта орналасуы керек.

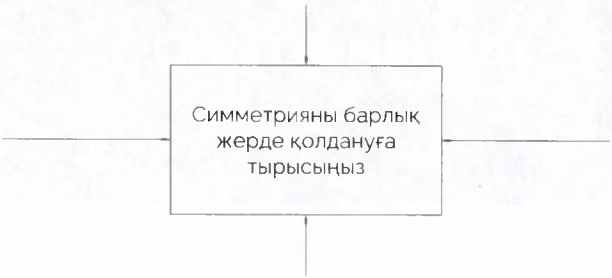


***Fig. 40**. Alignment of visual elements in preference to page borders*

16. Ғылыми этика

1 Әділдік

Барлық ғылыми коммуникацияларда адалдыққа, әділеттілікке ұмтылыңыз. Деректерді, нәтижелерді, әдістер мен процедураларды жариялауда адал және әділ болыңыз. Деректерді жасанды түрде қолдан құрмаңыз және деректерді заңсыз өзгертіп, бұрмаламаңыз. Әріптестеріңізді, зерттеу жұмысына көмек берген демеушілеріңізді немесе оқырмандарыңызды алдамаңыз.

Объективтілік 2

Эксперимент барысында деректерді талдау, одарды интерпретациялау, әріптестмен жұмыс кезінде шешім қабылдау барысында және басқа да объективтілікті талап ететін ғылыми-зерттеудің түрлі аспектілерінде біржақтылыққа жол бермеуге тырысыңыз. Біржақтылықтан немесе өзіңізді алдаудан аулақ болыңыз.

3 Тұтастық

Өз уәделеріңіз бен келісімдеріңізді орындауға тырысыңыз. Шын ниетпен әрекет етіп, ой мен іс-әрекеттің бірлігіне қол жеткізуге ұмтылыңыз.

Сақтық 4

Абайсызда қателіктер жібреру мен немқұрайлылықтан аулақ болыңыз. Өзіңіздің және әріптестеріңіздің жұмысын мұқият әрі сыни көзбен қарастырыңыз. Деректерді жинау, зерттеу және журнал қызметкерлерімен хат алмасу барысында әдептілік, дұрыс жазу жолдарын қолданыңыз.

5 Ашықтық

Деректерді, нәтижелерді, идеяларды, құралдарды, ресурстарды ашық түрде бөлісіңіз. Сынға және жаңа идеяларға ашықтық ниетте болыңыз.

6 Мөлдірлік

Зерттеуіңізді әділ бағалау үшін қажетті әдістерді, материалдарды, болжамдарды, талдауларды және басқа ақпаратты ашық түрде сақтаңыз.

Жауапкершілік 7

Зерттеу кезінде өзіңіз қатысқан бөліктің жұмысына үшін жауапкершілікпен қарап, зерттеу жобасында не істегеніңіз және не үшін істегеніңіз туралы түсініктеме немесе негіздеме түрінде есеп беруге дайын болыңыз.

8 Зияткерлік меншік

Патенттерді, авторлық құқықтарды және зияткерлік меншіктің басқа да нысандарын құрметтеңіз. Жарияланбаған деректерді, әдістерді немесе нәтижелерді рұқсатсыз пайдаланбаңыз. Зерттеуге қосқан үлестердің барлығына да тиісінше тиісті құрмет көрсетіңіз. Ешқашан да плагиат жасамаңыз.

9 Құпиялылық

Жеке жазбаларды, коммерциялық немесе әскери құпиялар мен емделушінің жазбалары сияқты маңызды деректерді құпия түрде қорғаныста ұстаңыз.

Жауапты басылым 10

Өз зерттеулеріңізді қызметтегі мансабыңызды өсіру үшін емес, зерттеулер мен стипендияны ілгері жылжыту үшін жариялаңыз. Ысырапшылдықтан және қайталанатын жарияланымдардан аулақ болыңыз.

11 Әріптестерге құрмет

Әріптестеріңізді құрметтеңіз және олармен әділ қарым-қатынаста болыңыз.

Әлеуметтік жауапкершілік 12

Зерттеу, қоғамдық білім беру және үгіт-насихат жұмыстары арқылы әлеуметтік игілікті алға жылжытуға және әлеуметке зиян келтірмеуге тырысыңыз.

13 Дискриминацияға жол бермеу

Әріптестер мен студенттерді жынысына, нәсіліне, ұлтына немесе ғылыми құзыреттілік пен адалдыққа қатысы жоқ басқа факторларға байланысты кемсітуге жол бермеңіз.

Біліктілік 14

Өмір бойы білім алу және оқу арқылы өзіңіздің кәсіби құзыреттілігіңіз бен тәжірибеңізді сақтау және жетілдіру керек. Ғылымдағы құзыреттілікті ілгерілету шараларын қабылдауыңызға тырысыңыз.

Заңдылық 15

Тиісті заңдар мен институционалдық, үкіметтік саясаттарды біліп және оларға бағынуға тырысуыңыз керек.

16 Жануарларға қамқорлық

Жануарларды зерттеу жұмысында пайдалану кезінде оларға лайықты құрмет пен қамқорлық көрсетіңіз. Жануарларға қажетсіз немесе нашар жобаланған эксперименттер жасаудан аулақ болыңыз.

Біліктілік 17

Өмір бойы білім алу және оқу арқылы өзіңіздің кәсіби құзыреттілігіңіз бен тәжірибеңізді сақтау және жетілдіруіңіз керек. Ғылымдағы құзыреттілікті ілгерілету шараларын қабылдауға тырысыңыз.

Адам субъектілерін қорғау 18

Адамдарға зерттеу жүргізген кезде оларға зиян келтіру мен тәуекелді барынша азайтып, пайда келтіруді барынша арттырыңыз. Адамның қадір-қасиетіне, жеке өміріне қол сұқпаңыз. Халық арасындағы осал топтарға арнайы сақтық шараларын қолдануды қабылдаңыз және зерттеудің пайдасы мен ауыртпалығын әділдікпен бөлуге тырысыңыз.

16. Конференция мен журналды таңдау

Зерттеу-мақаласын жазып болған соң, оны журналға немесе конференцияда жариялау керек. Алайда, мақаланы жазбас бұрын конференция үшін немесе журналға жіберуге мақала жазатынын анықтап алу керек. Егер журналға жіберуге мақала жазсаңыз, онда оның көлемі көбірек болғаны дұрыс (20 беттей). Ал егер мақаланы конференцияға жіберетін болсаңыз, онда оның бет саны азырақ болады (4—12 бет). Журналды немесе конференцияны таңдау барысында кездейсоқ немесе жақын арада болатын конференцияны таңдау дұрыс болмайды. Басты критерия — конференция немесе журнал сіз жазып жатқан мақалаға барынша сәйкес келетін болуы керек. Өз саланыздағы конференция мен журналдарға шолу жасап, мақаланы олардың ішіндегі ең танымалдарына мақала жіберуге тырысыңыз. Жұмысыңыздың сапасына қарай оның қабылдану ықтималдылығы да түрліше болады.

Кітап дизайнері: Мағжан **ҚАЙРАНБАЙ**

Кітап редакторы: Жұмабай **ҚАЙРАНБАЙ**

Кітап дизайнері, авторы
Мағжан Қайранбай

Редакторы
Жұмабай Қайранбай

ЖАСАНДЫ ИНТЕЛЕКТ САЛАСЫ БОЙЫНША МАҚАЛА ЖАЗУ ҚҰРАЛЫ

Басуға 22.09.2023 жылы қол қойылды.
Пішімі 60x84/16.
Көлемі 8,1 б.т.
Тапсырыс № 0028468.
Таралымы 500 дана.

ЖШС «Lantar Books»
Тел.: +7 702 251 02 17
e-mail: lantar2018@mail.ru
Алматы қаласы.


SDU
University Press


LANTAR BOOKS
ИЗДАТЕЛЬСТВО
"ЛАНТАР BOOKS"